Winter 1994

A theoretical study of the semiconductor laser structures with lateral discontinuity in the optical cavity

Yi Cai
New Jersey Institute of Technology

ABSTRACT

A Theoretical Study of the Semiconductor Laser Structures with Lateral Discontinuity in the Optical Cavity

by
Yi Cai

A theoretical study of the semiconductor laser structures with lateral discontinuity in the optical resonant cavity is presented in this thesis. Specifically, the lateral discontinuity is referred to the lateral expansion structure newly invented for increasing the power output of semiconductor lasers while keeping single mode operation.

In the first part of the thesis (Chapter 2 and Chapter 3), an explicit expression for calculating the lateral discontinuity problems is formulated by the incorporation of mode-matching method. Our approach is based on the mode expansion theory developed for lossless micro-wave and optical fiber waveguides, but the effect of gain in active laser media is discussed. Direct numerical evaluation to various discontinuities problems in semiconductor laser stripes are feasible, as long as the discontinuity lies entirely in a surface perpendicular to the direction of propagation.

The second part of the thesis is devoted to applications of the theoretical approach. A computer program written in Fortran-5.0 is used to calculate the efficiency of lateral expansion laser structure. Graphs of the electric field patterns are produced by a program written in programming language C.

A THEORETICAL STUDY OF THE SEMICONDUCTOR LASER STRUCTURES WITH LATERAL DISCONTINUITY IN THE OPTICAL CAVITY

by
Yi Cai

A Thesis
Submitted to the Faculty of New Jersey
Institute of Technology in
Partial Fulfillment of the Requirements for the Degree of
Master of Science in Applied Physics

Department of Physics

January 1994

A THEORETICAL STUDY OF THE SEMICONDUCTOR LASER STRUCTURES WITH LATERAL DISCONTINUITY IN THE OPTICAL CAVITY

Yi Cai

Dr. Ken K. Chin, Thesis Advisor   Date
Professor of Physics, NJIT

Dr. John C. Hensel, Committee Member   Date
Distinguished Research Professor of Physics, NJIT

Dr. Yuan Yan, Committee Member   Date
Visiting Professor of Physics, NJIT

BIOGRAPHICAL SKETCH

Author: Yi Cai

Degree: Master of Science in Applied Physics

Date: January 1994

Undergraduate and Graduate Education:

- Master of Science in Applied Physics,
  New Jersey Institute of Technology, Newark, NJ, January, 1994

- Bachelor of Science in Electronics and Information System
  Beijing University, Beijing, People's Repblic of China, July, 1992

Major: Applied Physics

This thesis is dedicated to my wife
Qian Huang

ACKNOWLEDGMENT

I would like to express appreciation to Dr. Ken K. Chin for his inspiration, guidance and support during the entire course of this thesis. I have learned a lot from him during the past three semesters, not only knowledge, but also the scientific way of thinking.

Appreciation is also extended to the committee members, Dr. John C. Hensel, Dr.Yuan Yan for their serving on the examination committee and for evaluating this research.

Special thanks to Professor Guanhua Feng for his guidance. I would also like to thank my friends Hancheng Liang and Ganming Qin for their resourceful discussion on using several software. Finally, special thanks go to my wife Qian Huang for her help and encouragement in every aspect.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF FIGURES

LIST OF FIGURES
(Continued)

Figure **Page**

CHAPTER 1

INTRODUCTION

Recent advances in optical communications have simulated the sustained studies of semiconductor laser. Among them, the high power and single mode laser structures have been extensively investigated because they directly affect the performance of optical communication systems.

Up to date, a wide range of semiconductor structures has been studied. A practical high power single mode laser device is expected. But the conflict between obtaining high power output and keeping single mode operation is not yet solved. One of the important limiting factors to make a high power laser is the thermal stress exerted by the optical beam on the two mirrors of the semiconductor laser resonant cavity. To avoid this stress from damaging the mirrors, a larger light spot is required, so that the stress can be distributed on a wider area. A wide laser stripe may have a large light spot, but higher order transverse modes exist in this case. This is the dilemma: on one hand the high power output needs wide stripes, while on the other hand the single mode operation requires narrow ones. The research presented in this thesis is dedicated to solve the problem using lateral expansion structure invented and patented by Dr. Eugene I. Gorden.

The lateral expansion laser structure has not yet been experimentally tested. There are many controversies in this topic, perhaps partially because of the over simplified description given in the patent. No analytical or numerical data indicated the efficiency of this structure.

In this thesis, a general method is presented to treat transverse discontinuity problems in metal or dielectric waveguides and cavities. This numerical method is

based on the well-established mode matching method, which is developed for treating transverse discontinuities in a passive waveguide. The effects of active laser media, in addition to passive waveguide media, are discussed. We have developed a practical theoretical description and calculation approach dealing with semiconductor laser structure with a variable complex shaped stripe.

The second purpose of this thesis is to apply the model of calculation to a practical device structure. As a special case, the efficiency of the lateral expansion laser structure is demonstrated. The numerical results are presented and discussed in the second half of this thesis.

In my view, just as molecular beam epitaxy (MBE) technology enables a variety of quantum well or superlattice structure to be grown, electron beam lithography and other advanced technology enables countless choices in the shape of each semiconductor layer to be grown. There is no reason to keep the shape of the semiconductor laser stripe uniformly rectangular. Variation of the stripe shape may show interesting and technologically important new properties. A simple theory along with an accurate and practical mathematical approach plays an important role.

CHAPTER 2

TYPICAL SEMICONDUCTOR LASER
AND LATERAL EXPANSION LASER STRUCTURE

2.1 Typical Semiconductor Laser Structure

2.1.1 Introduction to Laser

The word LASER is an acronym for light amplification by stimulated emission of radiation, which sums up the operation of an important electronics, as well as optical device. The effort of creating a useful laser system involves the techniques of solid state (or gaseous) electronics, and the transmission and detection of laser signal make use of a broad range of electronics.

There are two kinds of radiation emissions, spontaneous emission and stimulated emission. The process of stimulated emission occurs randomly. The rate at which electrons fall from an upper level of energy E_2 to a lower level E_1 is at every instant proportional to the number of electrons remaining in E_2 (the population of E_2). If conditions are right, stimulated emission can make the decay time much shorter. The stimulus is provided by the presence of photons of the proper wavelength (Fig 2.1). The electron is induced to drop in energy from E_2 to E_1, contributing a photon whose wave is in phase with the radiation field. If this process continues and other electrons are stimulated to emit photons in the same fashion, a large radiation field can build up. That is how the laser operates and why its light is coherent.

The rate of stimulated emission is proportional to the instantaneous number of electrons in the upper level n_2 and to the energy density of the stimulating field $\rho(v_{12})$. B_{21} is the proportionality factor. The rate at which the electrons in E_1 absorb photons should also be proportional to $\rho(v_{12})$ and the electron population in

3

E_1, namely the a proportionality factor as B_{12}. Introducing still another coefficient A_{21}, we can write the rate of spontaneous emission as $A_{21}n_2$. For steady state the two emission rates must balance the rate of absorption to maintain constant populations n_1 and n_2 (Figure 2.2).

Photon field

E_2

E_1
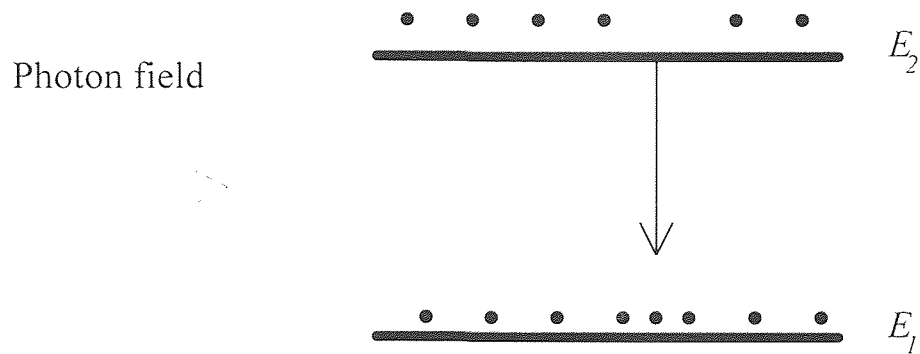
Figure 2.1 Stimulated transition of an electron from an upper state to a lower state, with accompanying photon emission

$$B_{12}n_1\rho(v_{12}) = A_{21}n_2 + B_{21}n_2\rho(v_{12}) \tag{2.1}$$

where $B_{21} = B_{12}$ according to quantum mechanics (principle of detailed balance).

At thermal equilibrium, the contribution of stimulated emission is negligible. The way to enhance the stimulated emission is to have a very large photon field energy density $\rho(v_{12})$. In the laser, this is encouraged by providing an optical resonant cavity in which the photon density can build up to a large value through multiple internal reflections at certain frequencies (v).

Similarly, to obtain more stimulated emission than absorption we must have $n_2 > n_1$:

$$n_2$$

Photon field

$B_{21}n_2\rho(v_{12})$ $B_{12}n_1\rho(v_{12})$ $A_{21}n_2$

E_2

$\rho(v_{12})$

E_1

$$n_1$$
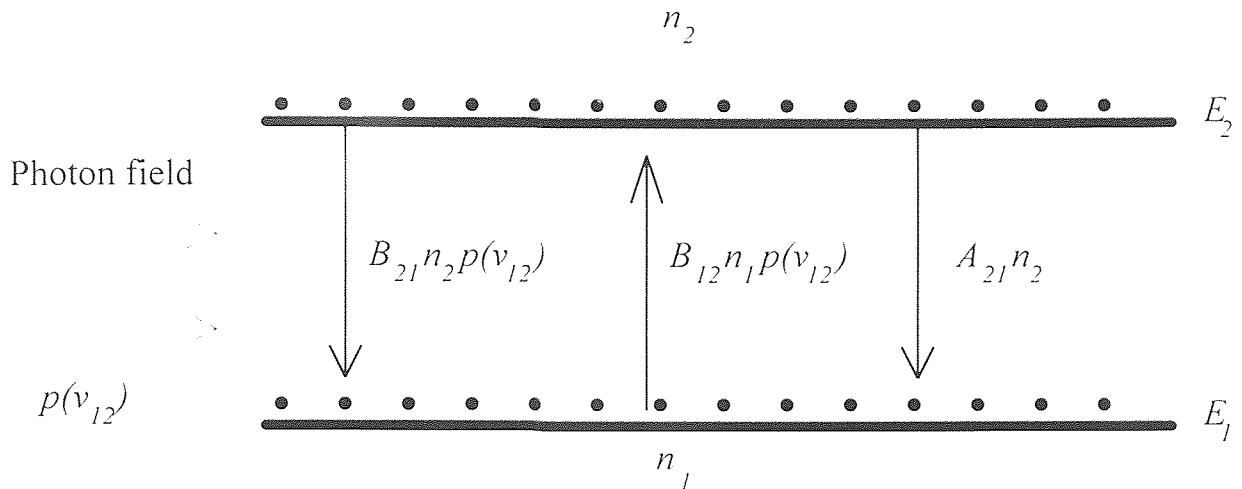
Figure 2.2 Balance of absorption and emission in steady state:
(a) Stimulated emission; (b) Absorption; (c) Spontaneous emission

$$\frac{\text{Stimulated emission rate}}{\text{Absorption rate}} = \frac{B_{21}n_2\rho(v_{12})}{B_{12}n_1\rho(v_{12})} = \frac{B_{21}}{B_{12}}\frac{n_2}{n_1} \tag{2.2}$$

This is a quite unnatural situation, called population inversion.

In summary, two fundamental requirements must be met to get lasing, an optical resonant cavity and population inversion.

2.1.2 Semiconductor Lasers

The laser became an important part of semiconductor device technology in 1962 when the first p-n junction lasers were built in GaAs (infrared) and GaAsP (visible). These devices differ from the solid, gas, and liquid lasers in several important respects. Junction lasers are remarkably small (typically on the order of $0.1 \times 0.1 \times 0.3$ mm), they exhibit high efficiency, and the laser output is easily modulated by controlling junction current. Semiconductor lasers operate at low pumping power compared, for example, with ruby or CO_2 laser. They compete with He-Ne lasers in power output. Thus the function of the semiconductor laser is

to provide a portable and easily controlled source of coherent radiation. Semiconductor lasers are particularly suitable for fiber optic communication systems.

The junction of the laser is formed between degenerate materials. The bands under forward bias are as shown in Figure 2.3. If the bias (and thus the current) is large enough, electrons and holes are injected into and across the transition region in considerable concentrations. As a result, the junction region is far from being depleted of carriers. The region contains large concentration of electrons within the conduction band and a large concentration of holes within the valence band. If these population densities are high enough, a condition of population inversion results, and the region about the junction over which it occurs is called an inversion region.

To build a p-n junction laser, we need to form a junction in a highly doped, direct semiconductor (GaAs, for example), construct a resonant cavity in the proper geometrical relationship to the junction, and make contact to the junction in a mounting which allows for efficient heat transfer. Beginning with a degenerate n-type sample, a heavily doped p region is formed on one side, for example by diffusing Zn into the n-type GaAs. At this point we have a large area planar p-n junction. Next, grooves are cut or etched along the length of the sample, leaving a series of long p regions isolated from each other. These p-n junctions can be cut or broken apart and then cleaved into devices of the desired length. At this point in the fabrication process, the very important requirement of a resonant cavity must be considered. It is necessary that the front and back faces be flat and parallel. This can be accomplished by cleaving. If the sample has been oriented so that the long junctions are perpendicular to a crystal plane of the material, it is possible to cleave the sample along this plane into laser devices, letting the crystal structure itself provides the parallel mirrors. The device is then mounted on a suitable

header, and contact is made to the p region. Various techniques are used to provide adequate heat sinking of the device for large forward current levels.
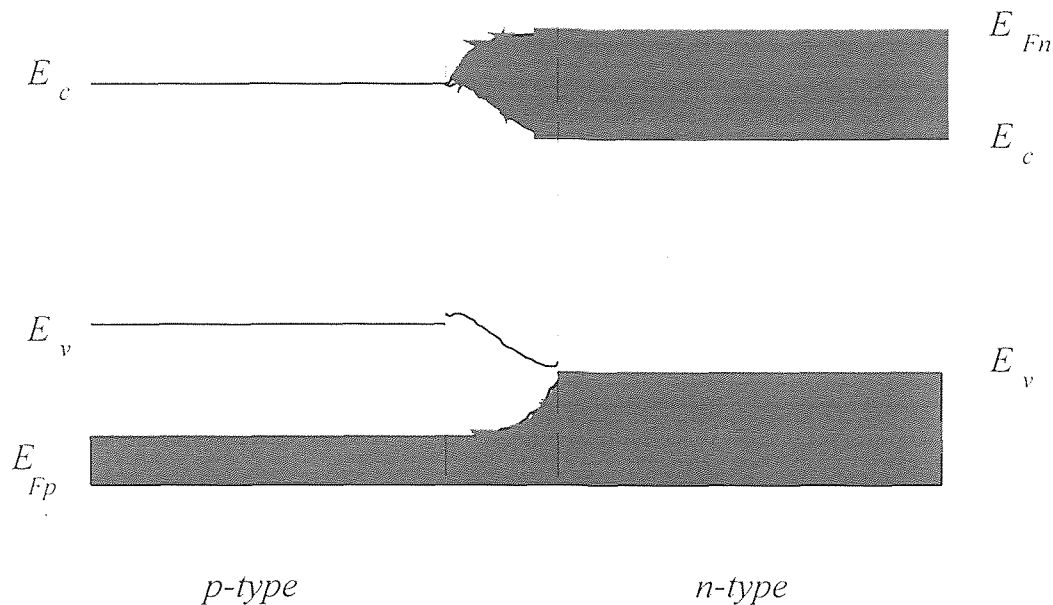
Figure 2.3 Band diagram of a p-n junction laser under forward bias.
(Region between dashed lines indicates inversion region.)

The above mentioned device is called homojunction laser. To obtain more efficient lasers, the active GaAs layer can be sandwiched between two AlGaAs layers Figure 2.4. This double-heterojunction structure improves the confinement of injected carriers to the active region. Furthermore, the change in refractive index at the GaAs-AlGaAs boundaries helps to confine the generated light waves. In the double-heterojunction laser shown in Figure 2.4b the injected current is restricted to a narrow stripe along the lasing direction, to reduce the total current required to drive the device. This type of laser was a major step forward in the development of laser for fiber-optic communications.

Proton bombarded region

Active layer stripe

-- p GaAs

-- p AlGaAs

-- undoped GaAs

-- n AlGaAs

-- n GaAs substrate

(a)   (b)
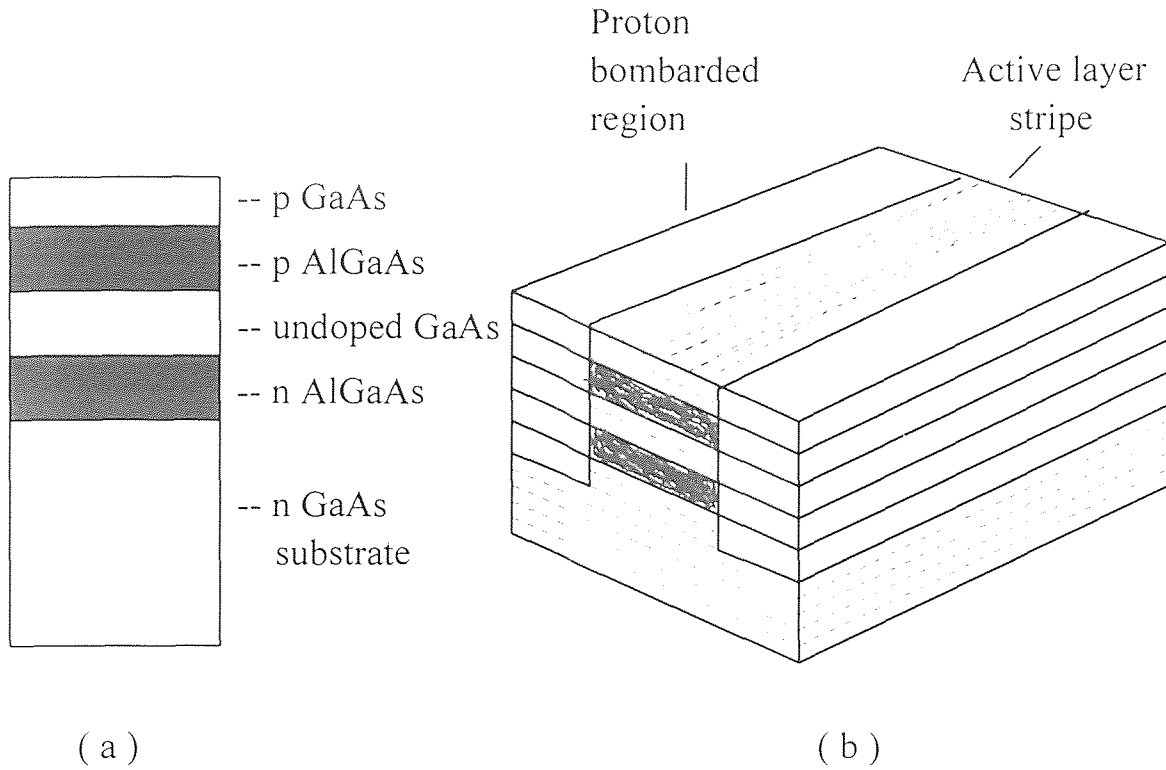
Figure 2.4 A double-heterojunction laser structure:

(a) Multiple layers used to confine injected carriers and provide waveguide for light;

(b) A stripe geometry designed to restrict the current injection to a narrow stripe along the lasing direction.

P

N   N

N

P   P
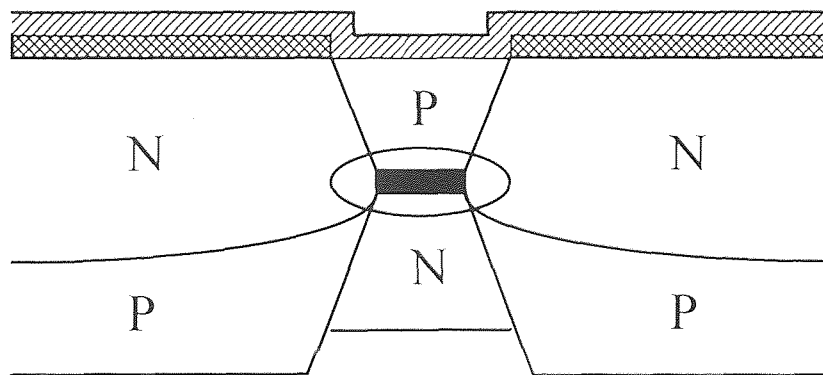
Figure 2.5 Schematic cross section of buried heterostructure index guided laser structure

In continual efforts to improve semiconductor laser performance, very complex structures requiring imaginative use of epitaxial growth have been employed. The graded index separate confinement heterostructure (GRINSCH), distributed feedback and buried heterostructure, as shown in Figure 2.5, lasers are invented.

2.2 The Limitation of typical laser structures

The initial application for semiconductor lasers was in fiber optical communications that required a single stable laser beam corresponding to lowest order mode operation. More recently, there has been an ever increasing demand for single mode lasers that can emit at significantly higher powers. Such devices could be used in optical recording, high speed printing, data distribution systems, analog signal transmission, long distance optical communication systems at high bit rates, and as pumps for solid state lasers. For all of these applications, it is desirable to have a high power laser that stable and reliably produces a single well-defined beam, i.e., lowest order mode output.

The most important type of semiconductor diode laser is the double hetero-structure (DH), as described in the previous section. The index of refraction of active layer is larger than the index of refraction of the surrounding cladding layers. Thus, the emitted radiation is transversely confined in a one dimensional dielectric waveguide formed by the two cladding layers and the active layer. For state of the art devices with active layers thinner than about 0.3 micrometers, the dielectric waveguide is such that only the fundamental transverse mode is supported on the direction perpendicular to the plane of the layers comprising the laser.

While the light is guided in the lowest order mode in the vertically transverse direction, such is not normally the case in the lateral direction (i.e., in

the plane of the layers). If a wide stripe contact is used to inject the pumping current, the optical output exhibits unstable, multimode and filamentary behavior, such as was the case with early diode lasers. The light versus current characteristic is highly nonlinear. The unstable filamentary behavior is exaggerated as one goes to higher and higher powers. Thus, such wide stripe structures, although they can produce the desired high power, have heretofore proven unsuitable for use in the typical applications contemplated for high power diode lasers, which applications require a single, stable, lowest order mode optical beam. However, they have found use where the requirement is simply infrared illumination.

Various techniques have been developed to provide for lateral confinement of emitted radiation, so as to achieve stable and reliable laser operation in the fundamental lateral mode. The simplest technique involves use of a narrow stripe contact on the upper surface of the laser. If the electric contact is shaped into a narrow stripe (less than 8 μm wide) running the length of the diode between the facets, the profile of the injected carriers established a weak waveguide that provides a type of current dependent guiding commonly referred to as gain guiding. However, these, structures exhibit strongly unstable multiple beam behavior at high powers and a non-linear light output versus current behavior known as kinking.

By introducing dielectric waveguide structures in the lateral direction as well as the vertically transverse direction, one gets index guided lasers. Using index guiding techniques and thin active layers, it is possible to produce a laser that supports only the fundamental transverse mode. Such lasers, having a two dimensional waveguide structure, emit a single mode, spatially coherent beam of light whose intensity profile at the facet is a bell-shaped surface. The output power is highly linear with current.

One type of index-guided laser is known as Buried Hetero-structure. In this structure a stripe-shaped relatively high index active region (e.g. GaAs) is surrounded vertically and laterally in transverse direction by lower index of refraction material (e.g. AlGaAs). Proper choice of dimensions assures lowest order mode operation. Other index guided structures, such as ridge guided structure, are also used.

However, the output power of such index guided laser formed using the AlGaAs-GaAs material system is limited by damage to the partially reflecting end facets of the diode laser. If an AlGaAs diode emits continuous wave optical power densities in excess of 6 to 9 mW per square micrometer of emitting area at the facet, the internal laser power density becomes so high that chemical reactions occur at the partially reflective end facets from which the light emerges. Stress is created and the end facet regions gradually darken, absorbing light, and the laser performance degrades. The output power also exhibits rapid time variations associated with the dynamics of the absorption process.

In addition, laser light is absorbed because of the nonradiative recombination of carriers at the end facets, where the boundary of the semiconductor material has a high density of surface states. At high optical power densities (20-25 mW per square micrometer) heavy radiation absorption at the facets introduces a thermal runaway process, which causes the mirrors to melt, thus causing catastrophic failure of the diode laser. The output power limits imposed by the gradual or catastrophic failure of the laser end facets have heretofore been extended by any of three known techniques.

First, the size of the lasing spot can be increased both perpendiculars to and parallel to the plane of the junction, to spread the emitted power over a wider area, thus allowing operation at higher power before the emitting facets are gradually degraded or catastrophically damaged. Vertical transverse spreading is limited in a

double heterojunction because a thin active layer is needed to achieve fundamental transverse mode behavior and there are practical limits, such as lifetime of the carriers, how thin the layer can be, etc. The thickness also affects the threshold current of the diode. Furthermore, while lateral spreading may lead to substantially increased power, it may also lead to the appearance of higher order modes. Thus, merely increasing the area of the laser spot, for example, by increasing the width of the active region in a stripe shaped structure will not lead to the achievement of a high power laser diode that operates in the fundamental lateral mode.

The second technique for boosting the power capability of a diode laser is to apply an anti-reflection coating to the front facet and a reflection coating to the back facet of the laser to increase the ratio of the laser's emitted power to internal power. This technique has limited value since in a typical laser the internal power is transmitted through the facets, even without an AR coating.

A third technique for increasing the power output of a laser is to prevent the mirror region from absorbing laser light or experiencing non-radiative recombination. Such structure is known as non-absorbing mirror or NAM structures. These are effective in increasing the catastrophic damage limit but are less efficient than conventional lasers.

The aforementioned structures have proved to be of limited usefulness in achieving reliable fundamental transverse and fundamental lateral mode behavior at high power.

A new structure, the so-called lateral expansion laser, is suggested, although not yet tested. It is expected to provide a high power semiconductor diode laser by increasing the lateral size of the laser spot while at the same time avoiding higher order lateral modes.

2.3 The Improvements in Lateral Expansion
Laser Structure

The lateral expansion laser diode is a stepped stripe double heterojunction laser that operates in the fundamental transverse mode. The wide portions of the stripe help to alleviate thermal stress, while the narrow portion of the stripe defines single mode operation. Vertically, fundamental transverse mode behavior is achieved by means of a conventional double heterostructure with a thin active layer. Fundamental lateral mode behavior in the wide stripe structure is achieved by suppressing the higher order lateral modes. Mode suppression is achieved by selectively increasing the loss of the higher order lateral modes.

Figure 2.6 and Figure 2.7 show the embodiments of the lateral expansion structure. For purpose of clarity, the drawings have not been drawn to scale.
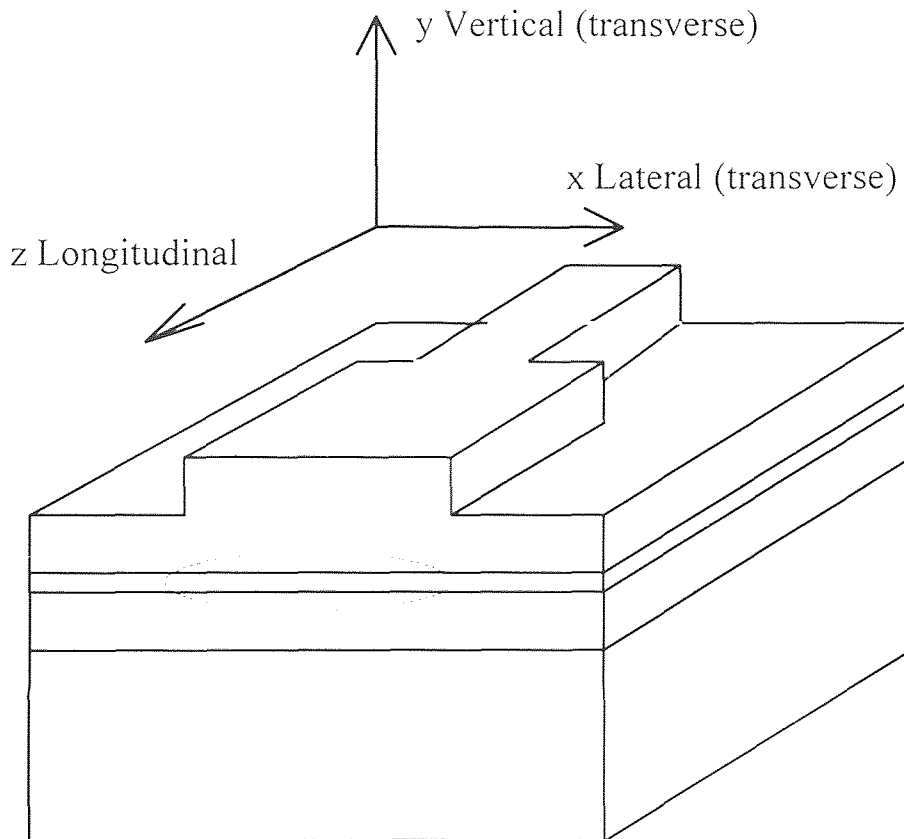
Figure 2.6 Lateral Expansion Structure
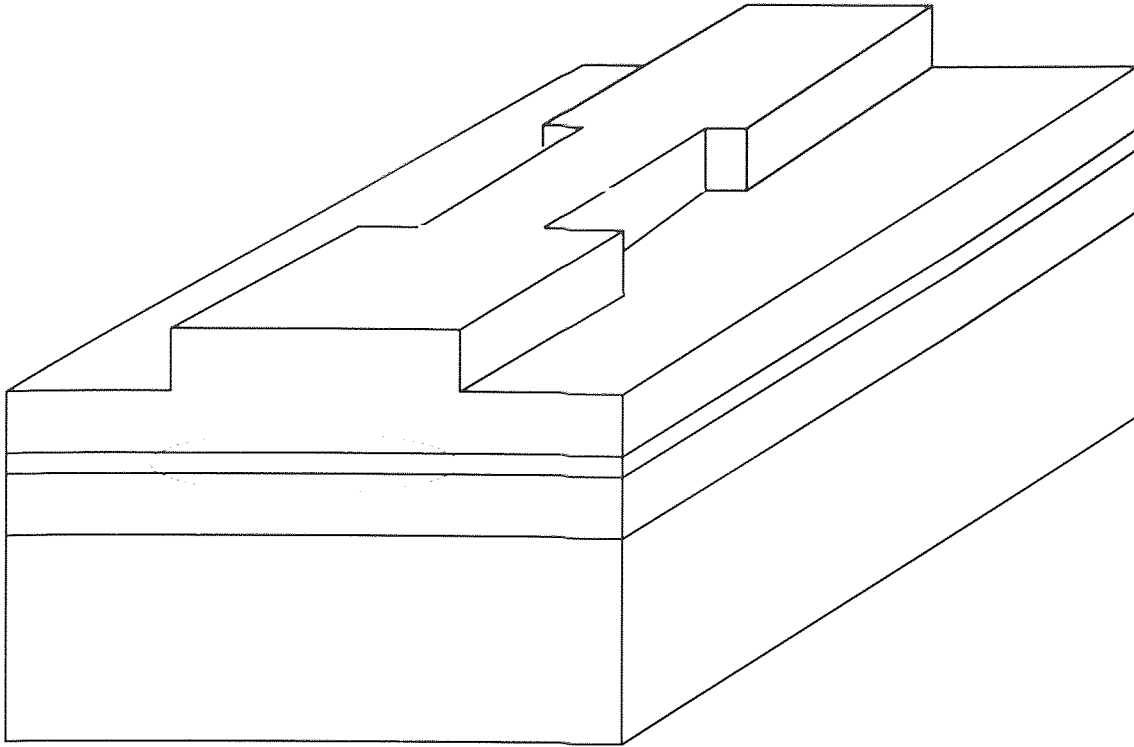
Figure 2.7 Lateral Expansion Structure (two ends expansion)

There are several structures that may be used to achieve fundamental lateral mode behavior in a wide stripe double heterojunction laser. One such structure utilizes the mode dependent coupling loss that occurs when a wide laterally wave-guided active region is formed longitudinally adjacent and contiguous with a narrow laterally wave-guided active region. The wide wave-guided region is normally wide enough to support higher order lateral modes while the narrow wave-guided region can support only the fundamental lateral mode. However, the coupling between the narrow fundamental and wide multimode wave-guided region is such that the coupling loss increases with increasing mode number of the multimode guide. The lowest order later mode has the smallest coupling loss and thus the lowest threshold current for lasing. Consequently, stable operation in the fundamental lateral mode is achieved in the wide waveguide. In other words, the wide waveguide region provides the high output power while the narrow region

provides the mode selectivity. Illustratively, the wide and narrow waveguide regions may be formed as ridged guided structures, with a wide ridge region defining the wide waveguide and a contiguous longitudinally adjacent narrow ridge defining the narrow waveguide.

According to the previous study done by Gorden, the suppression of higher order lateral modes as a result of coupling wide and narrow waveguide regions may be understood as follows. The power reflectivity at the interface between the wide and narrow waveguide is

$$R_m = C_{mp}^2 \, G^2 \, R \qquad (2.3)$$

where R_m is the effective reflectivity at the non-faceted end of the wide waveguide, R is the reflectivity at the facet, C_{mp} is the coupling coefficient for the mth order lateral mode of the wide waveguide to the fundamental mode of the narrow waveguide and G is the net gain of the active layer inside the narrow and wide waveguides. Note that this gain may be larger than or smaller than 1, in order to get any kind of lasing operation at all. C_{mp} is given by the following equation:

$$C_{mp} = (W_2/W_1)(1 - m^2 W_2^2/W_1^2) \qquad (2.4)$$

where W_1 is the width of the wide waveguide, and W_2 is the width of the narrow waveguide.

The equations indicate that the effective reflectivity goes down sharply as the mode number goes up and that the coupling loss is least for the m=0 order mode. Accordingly, the wide waveguide portion of the laser operates stably in fundamental lateral mode.

Note that only a symmetrical mode of the wide waveguide can couple to the symmetrical mode of the narrow guide since the narrow waveguide is centered on the wide waveguide, i.e., both waveguides share the same longitudinal axis. Accordingly, centering the narrow waveguide on the wide waveguide is desirable for providing mode selectivity.

Thus, a wide stripe double heterojunction laser has been disclosed which operates in the lowest order lateral mode. Problems relating to gradual degradation or catastrophic failure of the end facets or heating resulting from high current density, have been alleviated because the area of the lasing spot has been increased while at the same time higher order lateral modes have been suppressed.

Another structure for achieving fundamental lateral mode operation comprises a laterally wave-guided active region normally wide enough to support higher order lateral modes that are coupled to an even wider longitudinally adjacent optically active region without lateral waveguiding means. This second region is terminated by a facet. Light propagating from the guided region to the unguided region diffracts out into the unguided region and, after reflection off the end facet of the laser, does not couple perceptively back into the guided region. Since the higher order lateral modes diffract more, they experience greater loss. The lowest order mode suffers the least loss, thus enabling the guided region to operate stably in the lowest order lateral mode.

In both of the aforementioned structures, the lowest order lateral mode has the lowest current threshold for lasing since it has the lowest loss. When the laser is turned on, the circulating power for the lowest order mode grows more rapidly than the circulating power for the higher order modes and saturates the gain, effectively blocking oscillation in the higher order modes, even though the current may exceed the threshold for oscillation in the high order modes.

Lowest order lateral mode operation may also be achieved in a wide stripe structure when distributed feedback is used. In this case, facet reflection is eliminated and the distributed feedback can operate only for the lowest order lateral mode.

CHAPTER 3

COMPUTATIONAL MODEL
AND CALCULATION METHOD

The efficiency of the lateral expansion structure described in Chapter 2 is not transparent. The size of the light spot at the mirrors are not easy to be determined, because of the stripe discontinuities are involved. The discontinuity introduces the non linearity. Waves of each mode incident on the discontinuity junction not only is partially reflected and transmitted but also excites other modes. Some of them are propagating modes, while the others can be evanescent modes.

This problem can not be solved as simply as solving boundary condition problem in a uniform waveguide. A model-matching method introduced in this chapter is used to treat the discontinuities encountered in this structure.
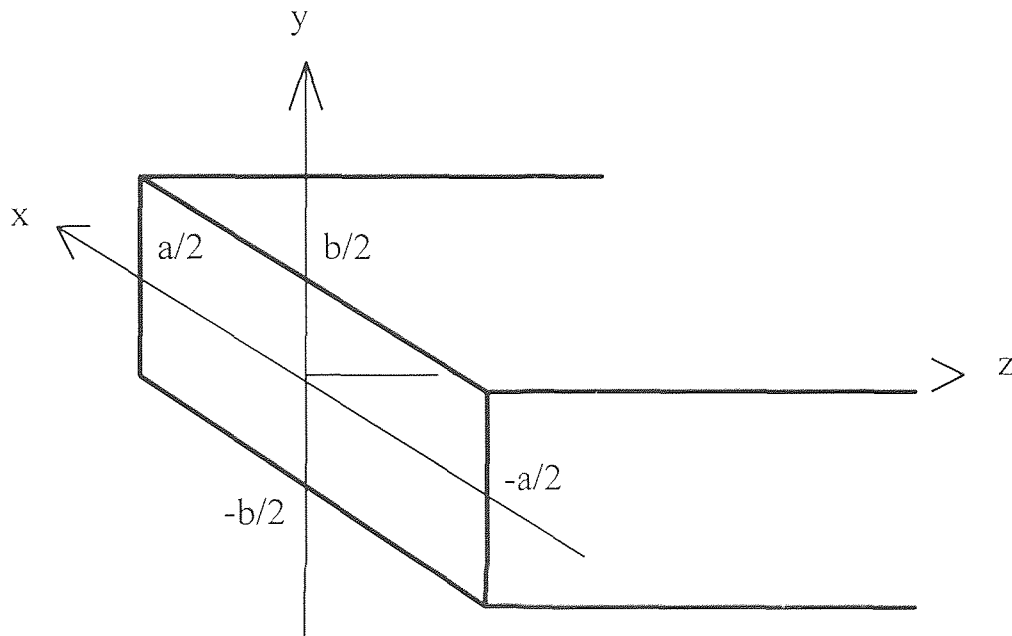
Figure 3.1 Rectangular metallic waveguide

3.1 Transverse Mode Functions

In the model-matching method, transverse electric and magnetic fields at the discontinuity plane are expanded in terms of normal modes of the two waveguides. It is necessary to know the mode functions in both waveguides.

The method has been tested first in rectangular metallic waveguides. The coordinates are as show in Figure 3.1. For a TM (transverse magnetic) mode solution, E_x and E_y are expressed in terms of E_z in the form of a gradient function

$$E_x = - \frac{\gamma}{\gamma^2 + \omega^2 \mu \varepsilon} \frac{\partial E_z}{\partial x} \tag{3.1}$$

$$E_y = - \frac{\gamma}{\gamma^2 + \omega^2 \mu \varepsilon} \frac{\partial E_z}{\partial y} \tag{3.2}$$

where γ is propagation coefficient; ω is angular frequency; μ is permeability; and ε is permittivity.

The longitudinal electric field E_z must satisfy the equation

$$\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_z}{\partial y^2} + (\gamma^2 + \omega^2 \mu \varepsilon) E_z = 0 \tag{3.3}$$

The boundary condition is that the tangential electric field shall vanish on the conductor.

The solutions in a rectangular waveguide of dimension a by b for the TM modes are as follows:

$$E_z = A \sin[m\pi(x-a/2)/a] \sin[n\pi(y-b/2)/b] e^{-\gamma z}$$

$$E_x = -j\beta A(m\pi/a)[(m\pi/a)^2+(n\pi/b)^2]^{-1} \cos[m\pi(x-a/2)/a]\sin[n\pi(y-b/2)/b]e^{-\gamma z}$$

$$E_y = -j\beta A(n\pi/b)[(m\pi/a)^2+(n\pi/b)^2]^{-1} \sin[m\pi(x-a/2)/a]\cos[n\pi(y-b/2)/b]e^{-\gamma z}$$

$$H_z = 0 \tag{3.4}$$

$$H_x = -\omega\mu E_y/\beta$$

$$H_y = \omega\mu E_x/\beta.$$

where β is also propagation constant, $j\beta = \gamma$; a and b are dimensions of the waveguide on x and y directions respectively. The integers, m and n, are mode numbers.

The propagation constant γ must be imaginary for a propagating wave, and must be real for a attenuated wave. The cut-off frequency is given by

$$f_{mn}=0.5v[(m\pi/a)^2+(n\pi/b)^2]^{-1/2} \tag{3.5}$$

where $v=1/(\mu\varepsilon)^{-1/2}$ is the velocity of plane waves in the unbounded medium.

Similarly, the solutions in a rectangular waveguide for the TE modes are as follows:

$$H_z = A\cos[m\pi(x-a/2)/a]\cos[n\pi(y-b/2)/b]e^{-\gamma z}$$

$$H_x = j\beta A(m\pi/a)[(m\pi/a)^2+(n\pi/b)^2]^{-1} \sin[m\pi(x-a/2)/a]\cos[n\pi(y-b/2)/b]e^{-\gamma z}$$

$$H_y = j\beta A(n\pi/b)[(m\pi/a)^2+(n\pi/b)^2]^{-1} \cos[m\pi(x-a/2)/a]\sin[n\pi(y-b/2)/b]e^{-\gamma z}$$

$$E_z = 0 \tag{3.6}$$

$$E_x = \omega\mu H_y/\beta$$

$$E_y = -\omega\mu H_x/\beta.$$

In the laser case, the thickness b is very thin compared with width a and wavelength λ. The eigen number n must be zero, if not the wave will be strongly attenuated. This can be seen readily from the following relations:

$$\gamma^2 = -[\omega^2\mu\varepsilon - (m\pi/a)^2 - (n\pi/b)^2] \tag{3.7}$$

If b is very small, the quantity in the brackets can become negative for any non zero n.

So, when dealing with dielectric waveguide case, y direction is trivial. The fields do not depend on y coordinate, when n is zero. The problem can be simplified as a dielectric slab guide. The dimensions are shown in Figure 3.2. Moreover, as mentioned at the end of the last chapter, only symmetric modes count, due the characteristics of the structure.
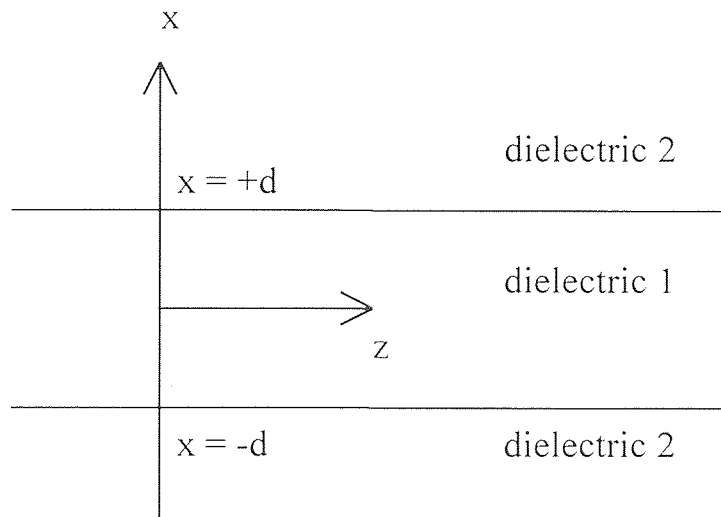
Figure 3.2 Dielectric slab geometry

This problem can be solved exactly, with no approximation in the wave equation for the vector potential. One can use solutions to the vector wave

equations of the electric field and magnetic field inside and outside the slab and matches the boundary conditions on the slab surface.

The symmetric solutions of dielectric slab guide for the TE modes are as follows:

$$E_y = A\cos(k_x x)e^{-\gamma z}, \quad |x| < d$$

$$E_y = Be^{-\alpha_x x}e^{-\gamma z}, \quad x > d$$

$$E_y = Be^{\alpha_x x}e^{-\gamma z}, \quad x > -d \tag{3.8}$$

$$H_x = (-\beta/\omega\mu_0)A\cos(k_x x)e^{-\gamma z}, \quad |x| < d$$

$$H_x = (-\beta/\omega\mu_0)Be^{-\alpha_x x}e^{-\gamma z}, \quad x > d$$

$$H_x = (-\beta/\omega\mu_0)Be^{\alpha_x x}e^{-\gamma z}, \quad x > -d$$

$$H_z = -j(k_x/\omega\mu_0)A\sin(k_x x)e^{-\gamma z}, \quad |x| < d$$

$$H_z = -j(\alpha_x/\omega\mu_0)Be^{-\alpha_x x}e^{-\gamma z}, \quad x > d$$

$$H_z = j(\alpha_x/\omega\mu_0)Be^{\alpha_x x}e^{-\gamma z}, \quad x > -d$$

where α_x is propagation constant in x direction. Taking advantage of symmetry, we need to match the boundary conditions only at x = d. Continuity of E_y/H_z at x = d gives

$$\tan k_x d = \alpha_x / k_x \tag{3.9}$$

We further have from the wave equation

$$\beta^2 - \alpha_x^2 = \omega^2\mu_0\varepsilon_2 \tag{3.10}$$

$$\beta^2 + k_x^2 = \omega^2\mu_0\varepsilon_1 \tag{3.11}$$

Combining these two equations, we find the determinantal equation

$$\frac{\alpha_x}{k_x} = \left[\frac{\omega^2 \mu_0 (\varepsilon - \varepsilon_i)}{k_x{}^2} - 1 \right]^{1/2} \tag{3.12}$$

As shown in Figure 3.3, the dispersion diagram, the dependence of the propagation constant β on frequency, is found by constructing an intersection of the functions *tan $k_x d$* and α_x / k_x for each frequency and mode and then evaluating β from the above equations. For decreasing ω, α_x / k_x moves toward the origin and intersections are lost, except for the intersection with the first branch of the tangent function. This corresponds to the dominant mode, m = 0, with no cutoff. All higher order modes have a cutoff; they are not guided below a certain critical frequency, the cutoff frequency.
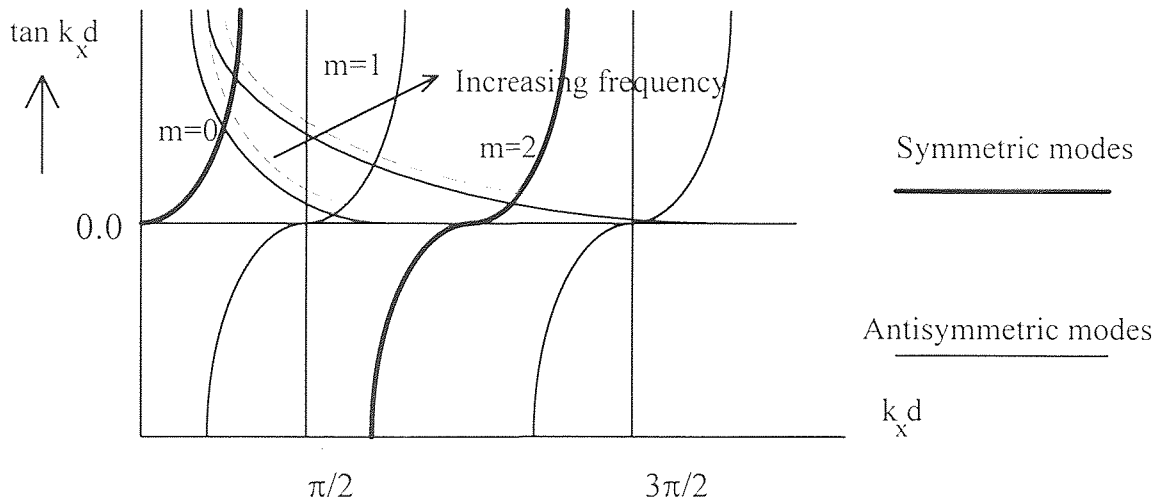
Figure 3.3 Graphical solution of the determinantal equation.
(Construction for TM modes shown dashed)

The symmetric solutions of dielectric slab guide for the TM modes are as follow:

$$H_y = A\cos(k_x x)e^{-\gamma z}, \quad |x| < d$$

$$H_y = Be^{-\alpha_x x}e^{-\gamma z}, \quad x > d$$

$$H_y = Be^{\alpha_x x}e^{-\gamma z}, \quad x < -d \qquad (3.13)$$

$$E_z = -(k_x/j\omega\varepsilon_i)A\sin(k_x x)e^{-\gamma z}, \quad |x| < d$$

$$E_z = -(\alpha_x/j\omega\varepsilon)Be^{-\alpha_x x}e^{-\gamma z}, \quad x > d$$

$$E_z = -(\alpha_x/j\omega\varepsilon)Be^{\alpha_x x}e^{-\gamma z}, \quad x < -d$$

Again we may construct a graphical solution to the transcendental equation:

$$\tan\frac{\alpha_x}{k_x} = \frac{\varepsilon_i}{\varepsilon}\left[\frac{\omega^2\mu_0(\varepsilon_i-\varepsilon) - k_x^2}{k_x}\right]^{1/2} \qquad (3.14)$$

as shown by dashed curves in Figure 3.3.

3.2 Computational Model

Discontinuity problems in surface waveguides have been studied for many years, partly because of their relevance to the design of dielectric rod antennas. More recently, developments in the field of integrated optics and fiber optics have caused renewed interest in the behavior of a junction between two surface waveguides, as shown in Figure 3.4a. Such waveguides may have either circular or planar geometry and the central core is surrounded by a finite cladding of lower permittivity, as shown in Figure 3.4b. If the geometry is planar, and both two cores of the waveguide are made with the same material, and both claddings are made with the same materials, as shown in Figure 3.4c, the model apparently represents the discontinuity junction encountered in the lateral expansion laser stripes.
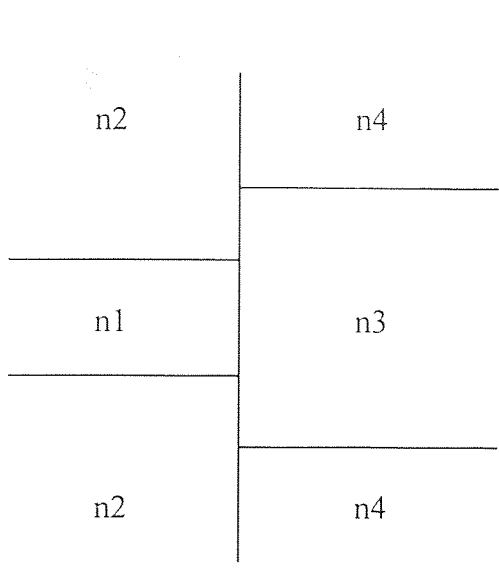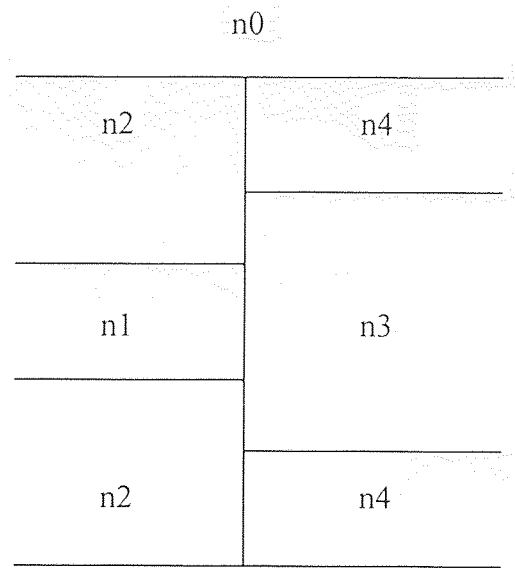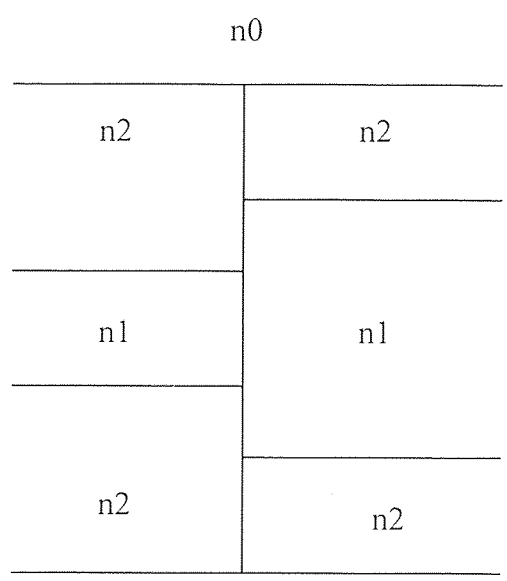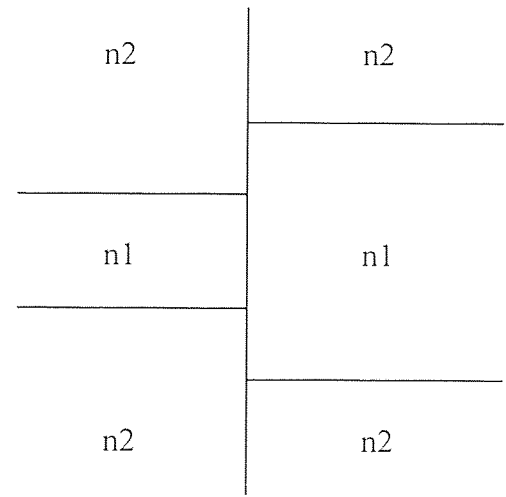
Figure 3.4 Discontinuities in surface waveguides

The analysis is simplified considerably if the dielectrics have infinite extent in transverse direction, as shown in Figure 3.4d. Although practical devices have finite transverse dimension, there is evidence, according to the previous study done by Anderson, Boyd and Mullen, to suggest that above assumption will not appreciably influence the excitation coefficients of the several surface waveguide modes with mode numbers are not too large. In metal waveguide case, there is no such problems, all possible modes can be included easily.

A different notation for TEmn mode is Hmn mode. As shown in Figure 3.5, when a fundamental mode wave incident on the junction from the left, waves of different modes are excited both in the left waveguide and the right one. The non linearity, brought by the discontinuity at the junction, is the cause of the excitation of all kinds of modes.
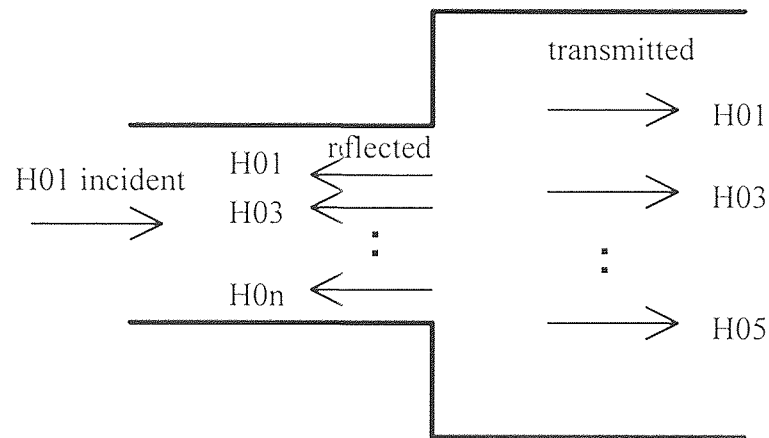
Figure 3.5 Modes excited at the discontinuity junction

The electromagnetic fields resulting from of the superposition of all the transmitted modes at the right side and of all the reflected modes plus incident mode at the left side must satisfy the boundary conditions at the junction plane. The boundary conditions are as follow:

$$n \times (H_2 - H_1) = 4\pi J/c$$

$$n \times (E_2 - E_1) = 0 \tag{3.15}$$

where J is surface current density. In the region where J is zero, as shown in Figure 3.6, and under the condition of fundamental mode incident, the boundary conditions become:

$$E_{y1} = E_{y2}$$

$$H_{x1} = H_{x2}$$

$$E_{x1} = E_{x2} = 0 \tag{3.16}$$

$$H_{y1} = H_{y2} = 0$$

(1)   (2)
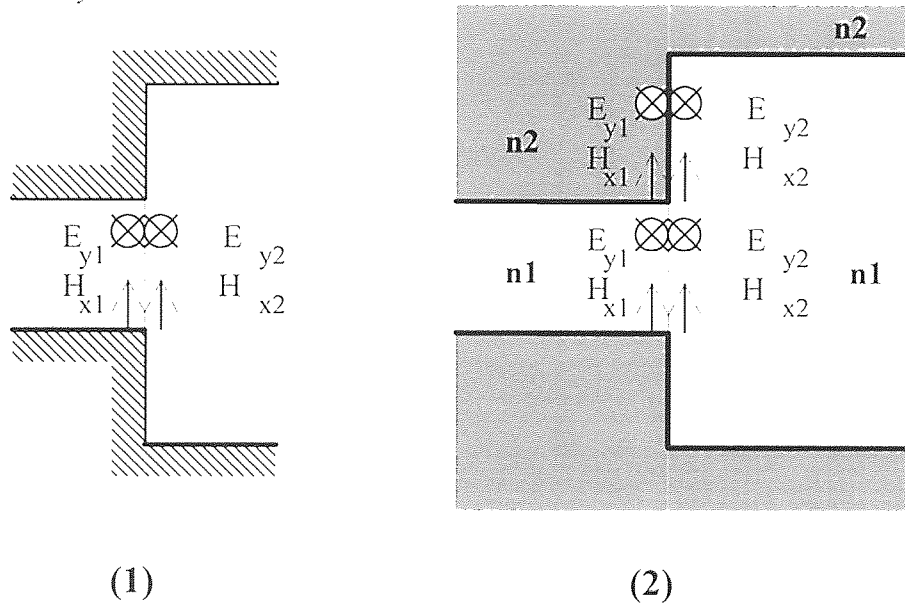
Figure 3.6 Boundary conditions in waveguide junction
(1). Metal waveguide; (2). Dielectric waveguide.

Although E_y and H_x are unknown functions, physically it is clear that they are constructed of all the possible waveguide transverse mode functions described in the previous section. In other words, E_y and H_x can be expanded in terms of transverse mode functions. To determine the amplitude or coefficient of each mode is the task for the mode matching method introduced in the next section.

3.3 Numerical Method for
the Solution of Waveguide Discontinuity Problems

A method for the solution of waveguide discontinuity problems using a digital computer is described in this section. For the method to be applied, the discontinuity must lie entirely in the surface transverse to the direction of propagation, and various techniques are described for treating planar discontinuity. Examples include a change in waveguide cross section and a change in the cross section of dielectric rods. Virtues of the method are simplicity and the ease with which discontinuity problems in overmoded waveguides may be treated.

List of principal symbols

a_{mp}, a_{Mp} = coefficient associate with mth or Mth mode

when pth mode is excited

a, d = widths of rectangular waveguides

e_t = transverse-electric-field mode function

E_t = transverse electric field

h_t = transverse-magnetic-field mode function

H_t = transverse magnetic field

m^*, M^* = maximum values of mode numbers m and M respectively

$$P_{mn} = \int_S e_{tn} \times h^*_{tm} \cdot dS = 0 \quad m \neq n$$

$$P_{mN} = \int_S e_{tm} \times h^*_{tN} \cdot dS = 0$$

$$P_{MN} = \int_{S'} e_{tM} \times h^*_{tN} \cdot dS = 0 \quad m \neq n$$

P_i = power associated with ith mode

$R = |R|\, e^{j\Phi}$ = complex reflection coefficient

$T = |T|\, e^{j\Psi}$ = complex transmission coefficient

γ_n, γ_N = propagation coefficients of nth and Nth modes, respectively

3.3.1 Introduction

In a previous study of discontinuities on shielded surface waveguides, Clarricoats and Slinn introduced a computer model-matching method to evaluate the transmission and reflection coefficients appropriate to a junction of circular dielectric rods. The method has been extended to an open planar-waveguide discontinuity, as shown in Figure 3.7b, by Clarricoats and Sharpe. Now it is going to be extended further to an optical cavity with active media, i.e. a laser structure.

The method described enables transverse discontinuity problems to be solved. It allows the complex reflection and transmission coefficients to be determined for modes incident on and excited at the junction, and, in principle, no limit is placed on the number of modes that can be treated. In practice, one seldom encounters situations in which more than a few modes can propagate in the waveguide that contains the discontinuity, and we have experienced no difficulty, as yet, in handling such problems. The method is such that the transverse cross-section of the waveguide or discontinuity may be chosen quite arbitrarily. However, so far we have only applied the method to waveguide and discontinuities of regular shape. Treatment of the arbitrary configuration requires the use of finite difference method, according to previous studies done by Hannaford, in order to obtain first the propagation coefficients and fields of all the significant modes incident on or excited at the junction.

3.3.2 Formulation of Basic Equations

The method now to be described applies to a metal waveguide junction of the kind shown in Figure 3.7a and, with a slight modification mentioned shortly, to those
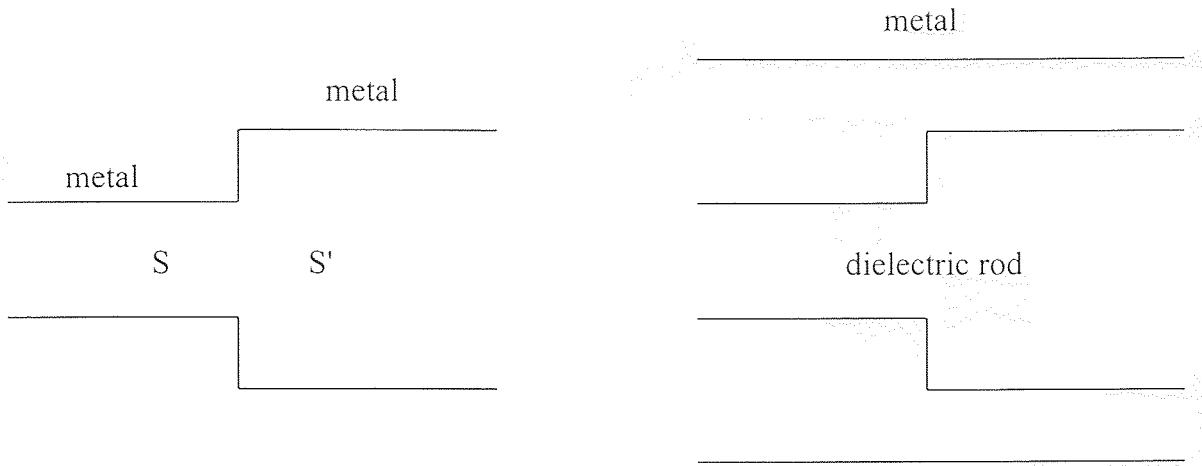
a) change in waveguide cross-section; b) change in dielectric rod cross-section
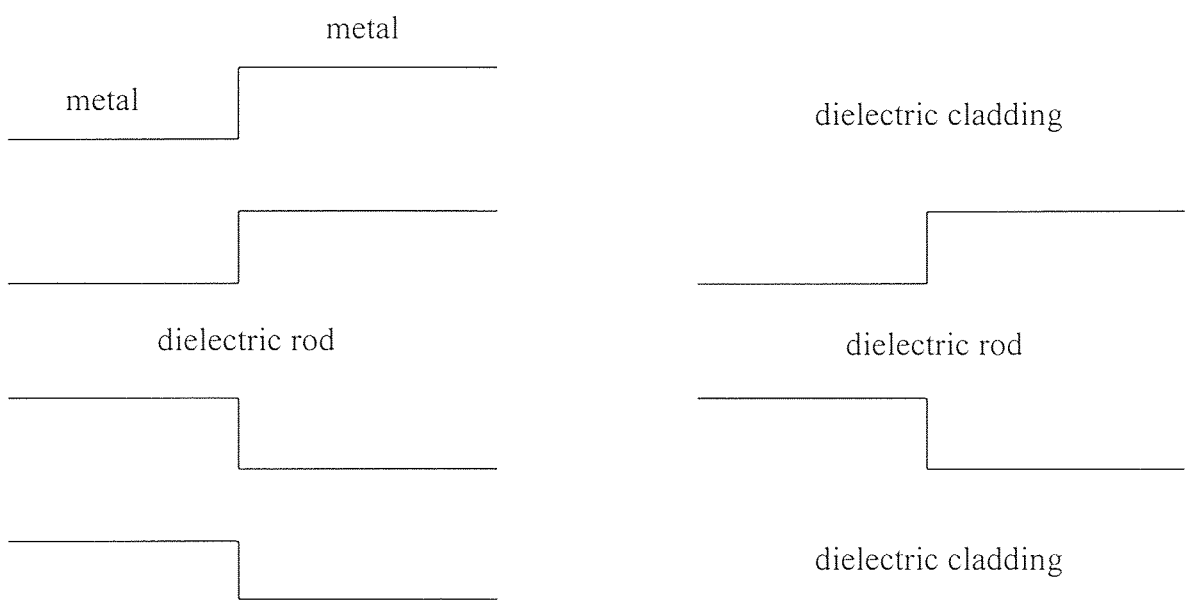
c) combination of junctions in a and b; d) dielectric waveguide change in cross-section

Figure 3.7 Transverse discontinuities

shown in Figure 3.7b and c. Then, the method can be extended to an open dielectric waveguide discontinuity shown in Figure 3.7d.

In essence, the transverse electric and transverse magnetic fields in the discontinuity aperture are expanded in terms of normal modes of the two waveguides. Two continuity equations are formed, and these are transformed into two sets of simultaneous equations suitable for the computation of the complex reflection and transmission coefficients of the normal modes.

If, in Figure 3.7a, the transverse electric field and transverse magnetic field in the aperture, of cross-section S, are E_t and H_t, respectively, and if e_{tm} and h_{tm}, e_{tM} and h_{tM} are corresponding transverse field functions for the mth normal modes in the two waveguides, assuming that the incident power is carried only by the pth mode in the left-hand waveguide:

$$Et = \sum_{m=1}^{\infty} (\delta_{mp} + a_{mp}) \, e_{tm} = \sum_{M=1}^{\infty} a_{Mp} \, e_{tM} \qquad (3.17)$$

$$Ht = \sum_{m=1}^{\infty} (\delta_{mp} - a_{mp}) \, h_{tm} = \sum_{M=1}^{\infty} a_{Mp} \, h_{tM} \qquad (3.18)$$

where $\delta_{mp} = 0 \qquad m \neq p$

$\qquad\qquad = 1 \qquad m = p$

e_{tm} and h_{tm} are chosen so that the incident power in the pth mode is given by

$$\text{Re} \int_S e_{tn} \times h^*_{tm} \cdot dS \, .$$

The complex coefficients a_{npp}, a_{Pp} then corresponds, respectively, to the reflection and transmission coefficients of the pth mode, and a_{mp}, a_{Mp} corresponds to the

transfer coefficients between the incident pth mode and the mth mode in the left-hand and right-hand waveguides, respectively. These coefficients are rather simply related to the elements of the generalized scattering matrix of the junction. We transform the equations to a form suitable for computation. Vector postmultiply equation 17 by h^*_{tm} and vector premultiply equation 18 by e^*_{tn}. Now, for metal waveguide, integrate over the aperture S and add to the right-hand side of equation 17 the term:

$$\sum_{m=1}^{\infty} \int_{S'-S} e_{tn} \times h^*_{tm} \cdot dS,$$

recognizing that this term is identically zero, since the total transverse electric field must vanish over the perfectly conducting transverse discontinuity. For the dielectric waveguide, all the integrals are over the whole junction plane. By using the well-known orthogonality relation stated in equation 21 and 23, equation 17 and 18 become:

$$\sum_{m=1}^{\infty} (\delta_{mp} + a_{mp}) P_{tm} = a_{Np} P_{NN} \tag{3.19}$$

$$(\delta_{np} - a_{np}) P^*_{nn} = \sum_{M=1}^{\infty} a_{Mp} P^*_{nM} \tag{3.20}$$

where

$$P_{mn} = \int_S e_{tn} \times h^*_{tm} \cdot dS = 0 \quad m \neq n \tag{3.21}$$

$$P_{mN} = \int_S e_{tm} \times h^*_{tN} \cdot dS \tag{3.22}$$

$$P_{MN} = \int_{S'} e_{tM} \times h^*_{tN} \cdot dS = 0 \quad m \neq n \tag{3.23}$$

If n takes integer values between 1 and ∞, two infinite sets of equations are obtained from which, in principle, all the coefficients a_{1p}, a_{2p}, ... a_{1p}, a_{2p}, ... may be obtained. In our approach to the solution of the discontinuity problem, the infinite series in these equations are truncated after a finite number of terms. Then, after evaluation of integrals of the form given in equations. 21-23, the two sets of equations may be solved simultaneously using a digital computer. Suppose that in the general equation 19 and 20, q terms and R terms, respectively, are taken in the infinite series. Equation 19 then defines a system of R terms, and equation 20 defines a system of q equations that may be combined in matrix form as set out below, equation (3.24).

The elements P_{ii}, P_{iJ} and P_{II} are either pure real or pure imaginary for lossless waveguides, which the coefficients a_{mp}, a_{Mp} are, in general, complex. Equation 8 may be subdivided into real and imaginary parts, so that the $q+R$ complex coefficients may be evaluated.

Physically, the truncation process is equivalent to assuming that only those evanescent modes up to the qth and Rth, respectively, contribute significantly to the stored energy at the discontinuity, a supposition that appears entirely reasonable for most practical cases. It was shown that, for a number of the discontinuities that have been studied, highly accurate results were obtained when only a few evanescent modes were included.

$$[P_1] [a] = [P_2] \tag{3.24}$$

$$[P_1] = \begin{bmatrix} P_{11} & P_{21} & \cdots & P_{q1} & \vdots & P_{11} & & & \\ P_{12} & P_{22} & \cdots & P_{q2} & \vdots & & P_{22} & & \\ \vdots & \vdots & \ddots & \vdots & \vdots & & & \ddots & \\ P_{1R} & P_{2R} & \cdots & P_{qR} & \vdots & & & & P_{RR} \\ \hdashline -P^*_{11} & & & & \vdots & P^*_{11} & P^*_{12} & \cdots & P^*_{1R} \\ & -P^*_{22} & & & \vdots & P^*_{21} & P^*_{22} & \cdots & P^*_{2R} \\ & & \ddots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ & & & -P^*_{qq} & \vdots & P^*_{q1} & P^*_{q2} & \cdots & P^*_{qR} \end{bmatrix}$$

$$[a] = \begin{bmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{qp} \\ \hdashline -a_{1p} \\ -a_{2p} \\ \vdots \\ -a_{Rp} \end{bmatrix} \qquad\qquad [P_2] = \begin{bmatrix} -P_{p1} \\ -P_{p2} \\ \vdots \\ -P_{pR} \\ \hdashline 0 \\ 0 \\ -P^*_{pp} \\ 0 \\ 0 \end{bmatrix}$$

This method may also be used to solve discontinuity problems in which the wave guide cross-section does not change, but the cross-section of a dielectric rod contained therein changes discontinuously in a plane transverse to the direction of propagation. A configuration of this kind is depicted in Figure 3.5b. In this case, the integrals all extended over the same wave guide cross-section S. When this method is used to solve an open dielectric discontinuity problem shown in Figure 3.5d, the integrals all extend over the infinite whole junction plane.

To summarize the method, the aperture field is expanded initially in the infinite sets of modes of the surface waveguides, and the transverse fields are matched across the junction plane. Assume a transverse-surface-wave mode incident from the left of the junction possessing unit amplitude. On application of well-known orthogonality conditions, two continuity equations are transformed into two infinite sets of equations containing the infinite sets of modal amplitudes a_n and a_M. To obtain an approximation solution to these equations, the series are truncated, and, in our study, we have included a total of nine modes on either side of the junction. A first step in the calculation of the coefficients is determination of the propagation coefficients and transverse fields of the modes, when the evaluation of the field integrals follows readily as they can be expressed in a close form.

3.4 Calculation

3.4.1 Solution of Eigenvalue Equation

The propagation constant k_x for the dielectric waveguide is the solution of eigenvalue equation:

$$\tan k_x d = \left[\frac{\omega^2 \mu_0 (\varepsilon - \varepsilon_i)}{k_x^2} - 1 \right]^{1/2}, \text{ for symmetric mode} \tag{3.25}$$

To find the solution by hand, one can plot the dispersion diagram and find the intersections.

A numerical method is used to get the solutions by computer, called midpoint method. Let us begin by considering a general equation of the form:

$$f(x) = 0 \tag{3.26}$$

Suppose it is known that this equation must have one and only one solution within the interval *[x0, x1]*, and *f(x0)<0* and *f(x1)>0*. If *f((x0+x1)/2)>0*, the root is now confined in a smaller interval *[x0, (x0+x1)/2]*. Otherwise, if *f((x0+x1)/2)>0*, the root must lies within *[(x0+x1)/2, x1]*. Then start all over again: find the value of *f(x)* at the midpoint of the interval; let its x value replace the left or the right end of the interval according to its sign of *f(x)*. Each time we repeat it, a smaller interval results. Till the interval is very small, the mid point of the interval is a good approximation of the root.

In our case the function is:

$$f(x) = \tan k_x d - \left[\frac{\omega^2 \mu_0 (\varepsilon - \varepsilon_i)}{k_x^2} - 1 \right]^{1/2}, \text{ for symmetric mode} \quad (3.28)$$

where $x = k_x d$.

There is a root in $[0, \pi/2]$, and one root in every π interval after $\pi/2$, until cutoff.

3.4.2 Numerical Integration

The elements of matrix [P1] and [P2] are results of the following integrals:

$$P_{mn} = \int_S e_{tm} \times h^*_{tn} \cdot dS$$

$$P_{mN} = \int_S e_{tm} \times h^*_{tN} \cdot dS \quad (3.29)$$

$$P_{MN} = \int_S e_{tM} \times h^*_{tN} \cdot dS$$

where

$$e_{tm} \times h^*_{tN} = e_{tmx} h^*_{tNy} k - e_{tmy} h^*_{tNx} k \quad (3.30)$$

As mentioned before the thickness of the active layer is so thin that only the fundamental mode exists. That is the mode number on the vertical direction is zero, then $h^*_{tNy} = 0$ and $e_{tmx} = 0$. Hence,

$$P_{mN} = \int_S e_{tm} \times h^*_{tN} \cdot dS = \int_S - e_{tmy} h^*_{tNx} dS \qquad (3.31)$$

Moreover, h^*_{tNx} and e_{tmy} are only the funtions of x. Then the surface integral can be replaced with a one dimension integral.

That is the problem of determining the area underneath a curve between two given values of the independent variables. Numerical integration method uses progressively smaller "panels" of measurable area to divide the area beneath the curve. The summation of the areas of these panels then provides an approximation of the total area. In the trapezoidal rule method, these panels are topped by straight-line secants to the curve, whereas Simpson's method tops each panel with a parabolic curve of its own. Both of these methods can be improved with an abbreviated Taylor series expansion to compute the error -- a method called end correction.

The method for numerical integration used in this thesis is the Simpson method with end correction. Since a parabola can be defined by a minimum of three points, we must divide the original area into an even number of panels. Each parabola is fitted to the tops of two adjacent panels. In general, we expect parabolas to produce a better fit than straight lines. That is, we should need fewer panels to obtain satisfactory convergence. The Simpson's rule formula with end correction looks like:

$$\left\{ 7(f_0 + f_n) + 14 \sum_{\substack{j=2 \\ j\ even}}^{n-2} f_j + 16 \sum_{\substack{j=1 \\ j\ odd}}^{n-1} f_j + \Delta x[f'(a) - f'(b)] \right\} \Delta x / 15 \qquad (3.32)$$

where

$f(x)$ is the function;

$f'(x)$ is the first order derivative of the function;

n is the number of panels;

integration is evaluated between the region of $x=a$ and $x=b$;

$\Delta x = (b - a) / n;$

$f_j = f(a + j \times \Delta x), \quad j \text{ is a integer.}$

The program automatically chooses the proper panel number n, to get results with satisfactory precision.

3.4.3 Simultaneous Equations with Complex Coefficients

The problem of determining the propagation coefficient for each mode, eventually becomes a problem of solving a set of simultaneous equations.

The technique solving the simultaneous equations used in this thesis is called Gauss elimination. With the Gauss method, the original equations are manipulated so that the coefficient matrix contains a value of unity at each point on the major diagonal and zero at each position below and to the left of the major diagonal.

Two basic types of matrix operations are used in the Gauss elimination method: scalar multiplication and addition. Any equation can be multiplied by a constant without changing the result. This is equivalent to multiplying one row of the coefficient matrix and the corresponding element in the constant vector by the same value. Also, any equation can be replaced by the sum of two equations.

After the coefficient matrix is trianglized, the very last equation can be solved since it has only one unknown. And the second last equation has two unknowns, but one of them is solved in the last equation. So, by substituting in the

solved unknowns in sequence, all the unknowns can be solved. The process is called back substitution.

But in the problem encountered in this thesis, all the elements in the simultaneous equations are complex numbers. It can not be treated easily by using the method described above. Modification to the method is needed.

First, n simultaneous equations with complex coefficients are converted into 2n equations with real coefficients. Then, the resulting equations can be solved by one of the methods developed previously in this section.

Consider a general statement of a set of two equations:

$$\begin{bmatrix} AR_{11} + jAI_{11} & AR_{12} + jAI_{21} \\ AR_{21} + jAI_{21} & AR_{22} + jAI_{22} \end{bmatrix} \begin{bmatrix} IR_1 + jII_1 \\ IR_1 + jII_1 \end{bmatrix} = \begin{bmatrix} VR_1 + jVI_1 \\ VR_2 + jVI_2 \end{bmatrix} \tag{3.33}$$

where the following symbols are used:

AR_{kl} = real part of coefficient k, l

AI_{kl} = imaginary part of coefficient k, l

IR_l = real part of unknown coefficient l

II_l = imaginary part of unknown coefficient l

VR_l = real part of constant l

VI_l = imaginary part of constant l

Multiplication of the terms on the left of the above equations produces groups that alternatively include the complex operator j. Separate the real part and imaginary part, which is equal respectively. The complete new equations can be summarized as equation (3.34).

Notice that each original coefficient appears twice in the new matrix. It can be proved that it is also true for the case of n larger than two. The solution vector

for the new set of equations can readily be found by the method of Gauss elimination.

$$
\begin{bmatrix}
AR_{11} & -AI_{11} & AR_{12} & -AI_{12} \\
AI_{11} & AR_{11} & AI_{12} & AR_{12} \\
AR_{21} & -AI_{21} & AR_{22} & -AI_{22} \\
AI_{21} & AR_{21} & AI_{22} & AR_{22}
\end{bmatrix}
\begin{bmatrix}
IR_1 \\
II_1 \\
IR_2 \\
II_2
\end{bmatrix}
=
\begin{bmatrix}
VR_1 \\
VI_1 \\
VR_2 \\
VI_2
\end{bmatrix}
\tag{3.34}
$$

The Gauss elimination method is not suitable for solving very large matrices. More and more multiplication and subtraction operations are performed as the number of equations increases. The resulting roundoff error can produce a meaningless solution. In that case Gauss-Seidel iterative method can be used. But in the problem discussed a maximum of 12 modes are considered, it has been demonstrated that Gauss elimination can give very precise results.

CHAPTER 4

ANALYSIS BY COMPUTER
SIMULATION METHOD

In these chapter results obtained by using numerical analysis are presented. The matchings of the field pattern at the junction plane are first discussed, since it indicates the satisfaction of the boundary conditions at the discontinuity junction. The almost perfect matches of the transverse electric fields at the junction indicate the the accuracy of this method. The convergence is the next topic. The electric field pattern away from the discontinuity junction in the propagation direction is also an important topic. It is closely related to the efficiency of the calculation of the lateral expansion laser structure.

The reason that makes the calculation results of dielectric waveguide cases less accurate than the metal waveguide case is the neglecting of the radiation modes. Only propagating modes are calculated in the dielectric case, because it is really complicated to solve the eigen value equation for the propagation constant k_x, while it is a complex variable. But as cited in the reference, neglecting the radiation modes can keep the k_x be real. In that case, even in principle, only a finite number of propagating modes can be used in the field expansion when using the method of mode matching. Since a set of finite number of functions can not form a complete set, thus result in the inaccuracy. It is less accurate when the width on one side is much larger than the width on the other side of the junction. That is when the radiation modes play an important role in matching the field pattern out of the narrow waveguide core.

A lot of discussion is based on the results obtained from metal waveguide problems, because of the results are relatively easier to obtain and more accurate.

Once a property of the structure is found in the metal waveguide, a comparison is made with the dielectric waveguide case. That is because to obtain results for dielectric waveguides are very time consuming, and less accurate. On the other hand metal waveguide is approximation for dielectric waveguide with very large refraction index difference between the cladding and core.

4.1. Matching the Field Profiles at the Discontinuity Junction

Figure 4.1 shows the matching of the transverse electric field profiles at the junction of metal waveguides. The waveguide width is 1um on the incident side, and 2um on the output side, with only fundamental mode incident. The profile of incedent mode is shown in Figure 4.2.
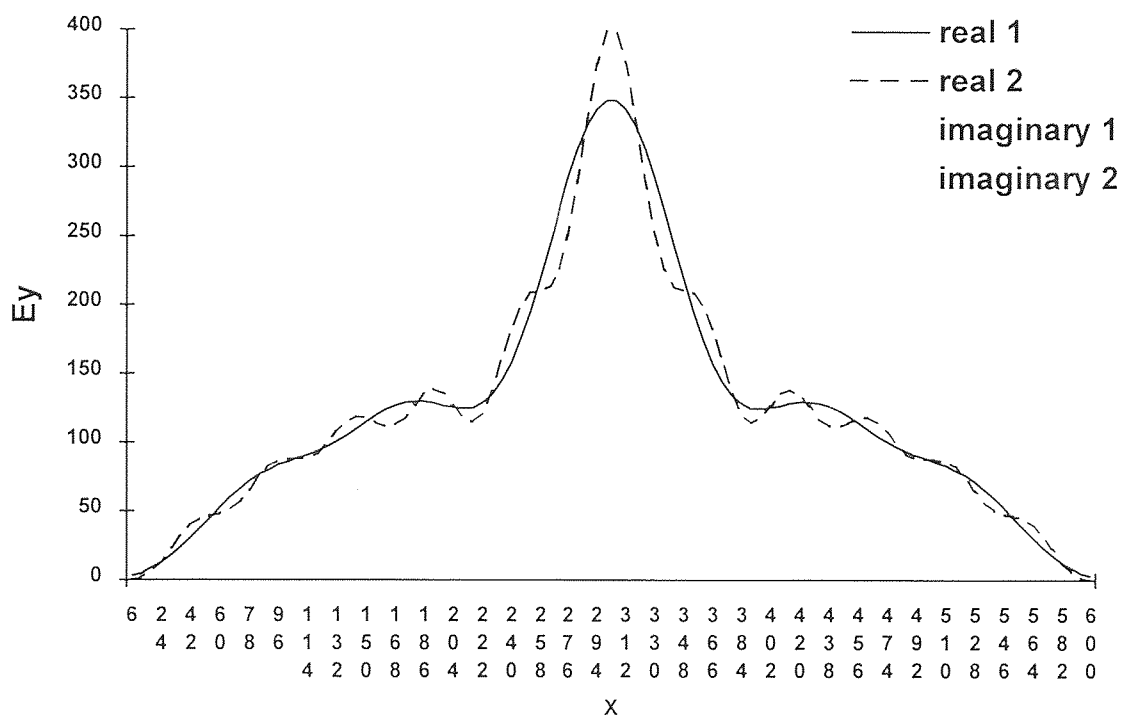
Figure 4.1 Electric field profiles matching at the junction.
(a=1um, b=2um, λ=1.68um, 12 mode approximation)
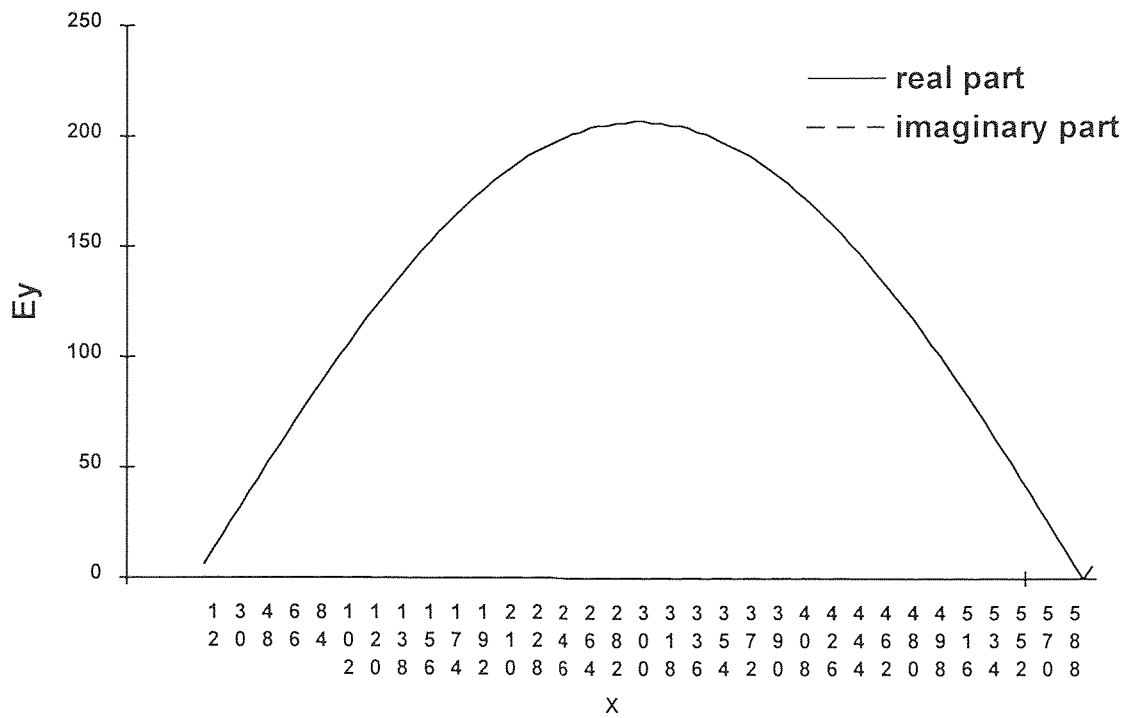
Figure 4.2   Electric field profile of H01 incident on the junction (a=1um, b=2um, λ=1.68um, 12 mode approximation)

from the 1um waveguide. Wave length is 1.68um. In that case the left side is a single mode waveguide. As shown in the figure, a 12 mode approximation made very accurate matching of the patterns in the aperture. These indicate the results obtained satisfy the boundary transverse electric field continuity condition. The transverse magnetic field H_x should also be continuous. As the electric field is physically measurable, only electric field is discussed.

Convergence requires that the more modes are used in the expansion, the more accurate result should be. A result obtained from six mode approximation is shown in Figure 4.3. Apparently, it is not as accurately matched as the twelve mode approximation results. These demonstrate the convergence.
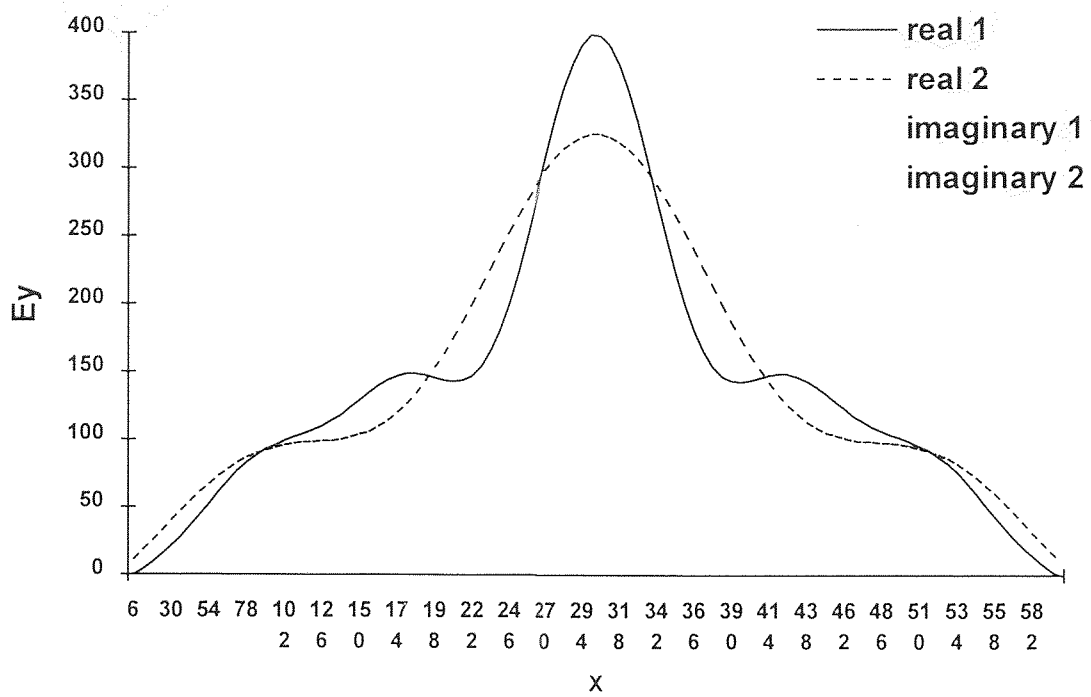
Figure 4.3   Electric field profiles matching at the junction.
(a=1um, b=2um, λ=1.68um, 6 mode approximation)
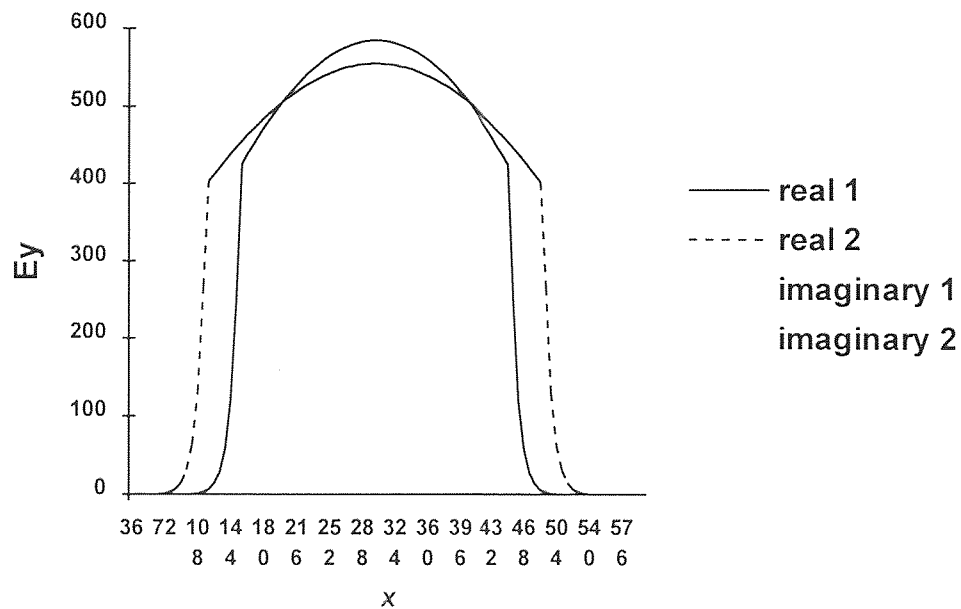
Figure 4.4   Electric field profiles matching at the junction.
(a=4um, b=5um, λ=0.88um, 5 mode approximation)
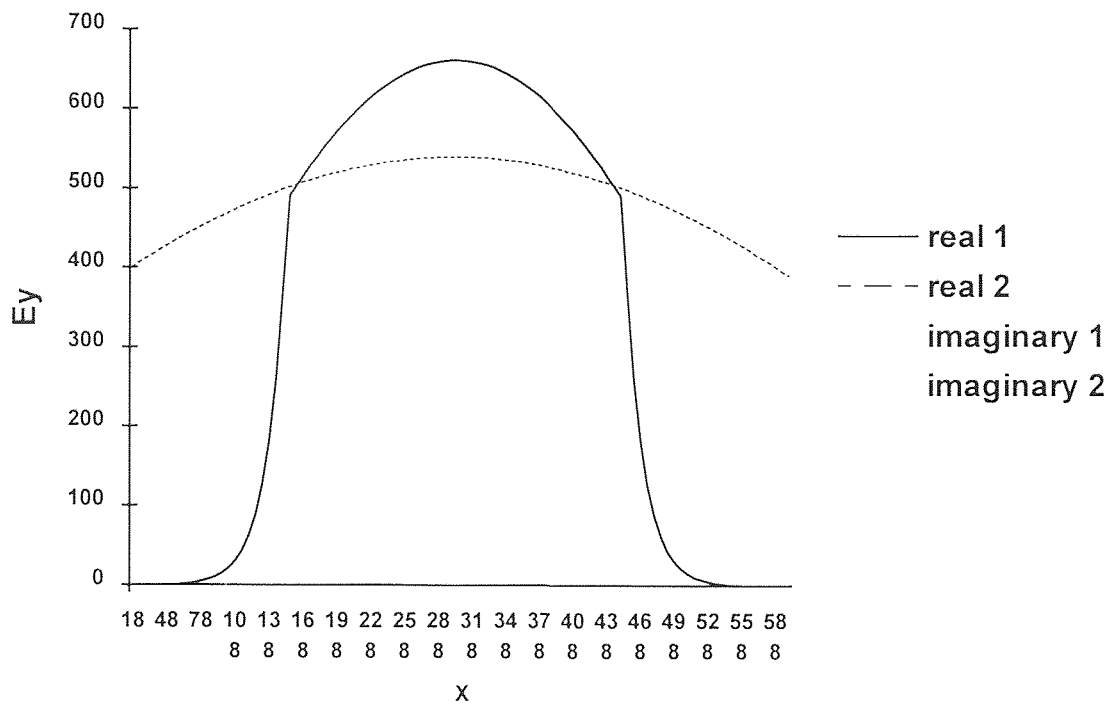
Figure 4.5 Electric field profiles matching at the junction.
(a=2um, b=4um, λ=0.88um, 3 mode approximation)

Similarly, results for dielectric waveguide discontinuity problems are shown in Figure 4.4 and Figure 4.5 for fewer modes and shorter wavelengths. Neglecting radiation modes made the results less accurate, especially when there is a large difference between an and b, as shown in Figure 4.5. The incedent mode profile is shown in Figure 4.6.

4.2. Mode Decomposition

The field patterns at the discontinuity junction, as shown in Figure 4.1 and Figure 4.3, are summations of all possible mode components. This is the basis of mode expansion. Convergence requires that the components of high order modes are gradually vanishing as mode number increases, otherwise there is no reason to truncate the mode expansion series.
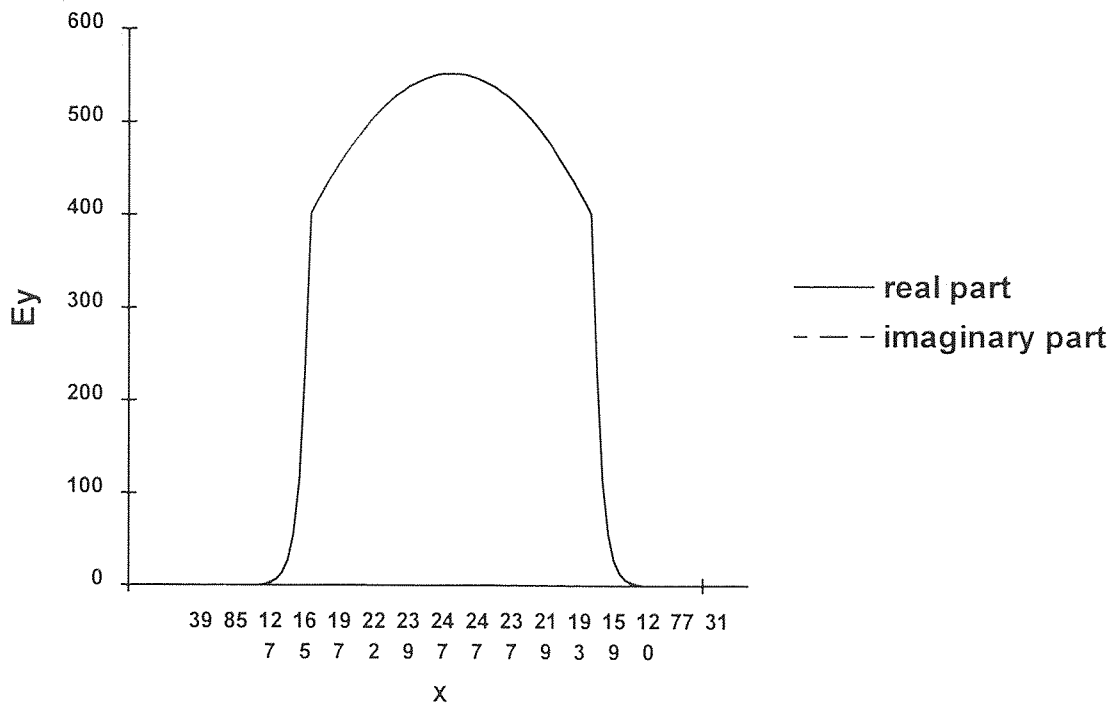
Figure 4.6   Electric field profile of H01 incident on the junction
(a=4um, b=5um, λ=0.88um, 5 mode approximation)
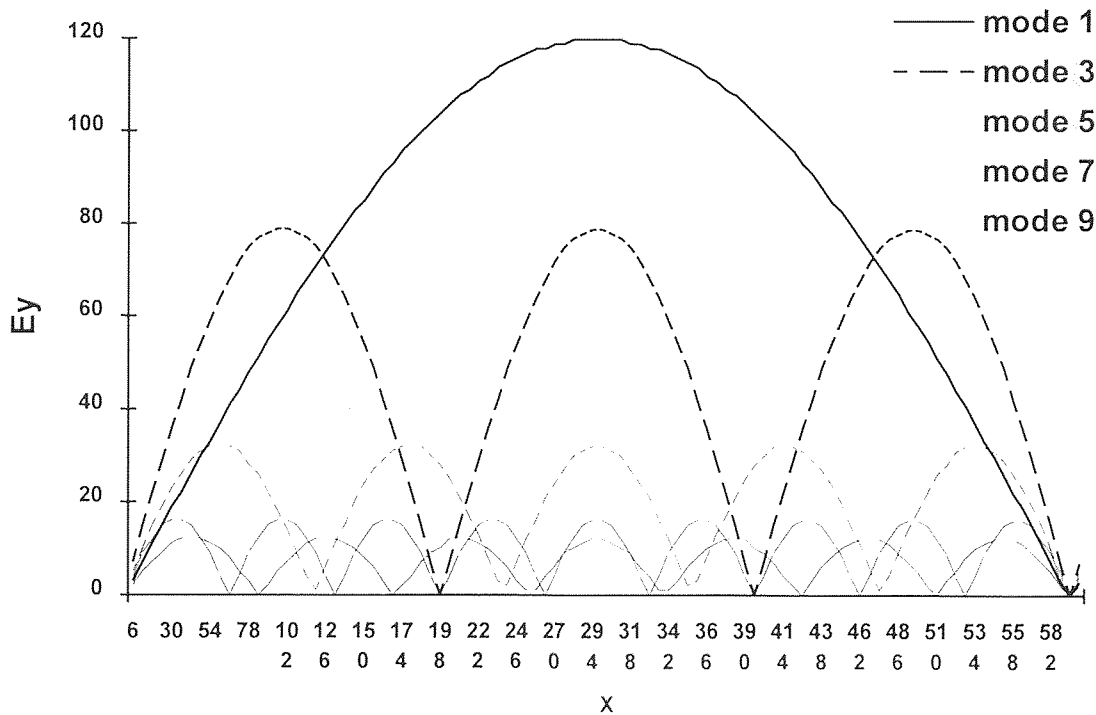
Figure 4.7   First five symmetric modes' components comparison
(a=1um, b=2um, λ=1.68um, 12 mode approximation)
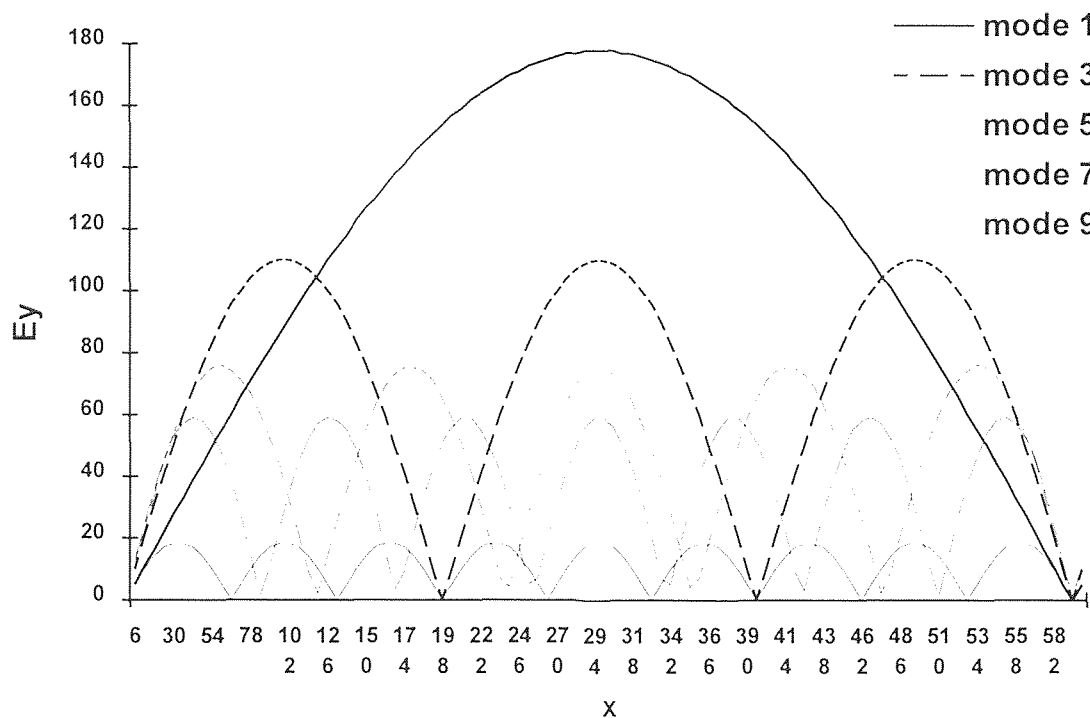
Figure 4.8 First five symmetric modes' components comparison
(a=1um, b=2um, λ=0.88um, 12 mode approximation)

Figure 4.7 and Figure 4.8 show the amplitudes of the first five symmetric modes on the out put side under two different wavelength wave incidents. The two figures are obtained from the same waveguide structure (incident side 1um of width, output side 2um of width, fundamental mode incident only). The only difference is the wave length.

Figure 4.7 shows the result when wave length is 1.68um. In the incident side only fundamental mode can propagate, in the output waveguide two modes can propagate, but among them only the fundamental mode is a symmetric mode, the other one is antisymmetric that can not be excited by the symmetric incident mode. This is the circumstance under which the lateral expansion structure is sufficient. It is going to be discussed in detail in the next section.

Figure 4.8 is obtained by using 0.88um light. In this case, the incident waveguide can propagate waves of two modes, among them one is symmetrical. The output waveguide can propagate waves of four modes, among them two are symmetrical.

Both Figure 4.7 and Figure 4.8 show the trend of convergence, as can be seen the low order modes dominate.

Although, their mode component looks similar to one another. The situation shown in Figure 4.8 is not sufficient in enlarging light spot as the one shown in Figure 4.7. This is the topic of the next section.

4.3 Electric Field Patterns at the Facet

The modes excited at the junction are classified as two types, propagating modes and evanescent modes. The propagate modes do not change their amplitude as they propagating in a lossless waveguide, while the evanescent modes exponentially decay as they propagate. So as they reach the facet, the field pattern is different from that of at the junction. The pattern at the facet is essentially comprised of only propagating modes.

For the time being, we assume there is a anti-reflection coating at the facet, so facet reflection is negligible. Precise solution needs to include the reflected wave. The reflected wave is partially reflected and partially transmitted by the junction when it heads back. But that is not the end of the story, the wave will go back and forth infinite times as shown in Figure 4.9. A precise field pattern can be obtained by adding up all these waves.
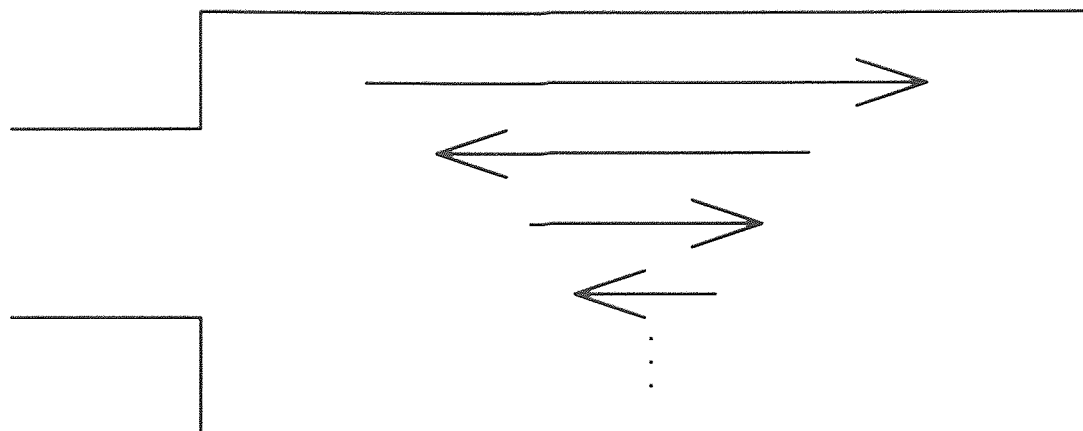
Figure 4.9 Wave reflected back and force between junction and facet

The electric field pattern changes from Figure 4.1 at the junction to Figure 4.10 at the facet 20um away from the junction. As shown in Figure 4.10, fundamental mode dominants in the wide waveguide that can propagate two modes. This shows the significance of the lateral expansion structure. The light spot is enlarged to twice as big as it is incident, thus thermal stress is distributed on a wider area on the facet. This can not be achieved by a simple shaped stripe 2um of width, because, in that case, one antisymmetric higher order mode will compete with the fundamental mode, resulting in a multimode laser.

The result for dielectric case is showed in Figure 4.12.

Figure 4.11 is obtained from Figure 4.3 when it is 20um away from the junction. Because the existence of higher order propagating mode in the output waveguide, the field patterns remains a sharp peak. The light spot does not spread out much. It is constructed by two propagating modes (mode 1 and mode 3). A higher order mode (mode 3) has a significant portion in the field pattern at the facet. That results in multimode property. Under this circumstance, the lateral expansion structure helps nothing.
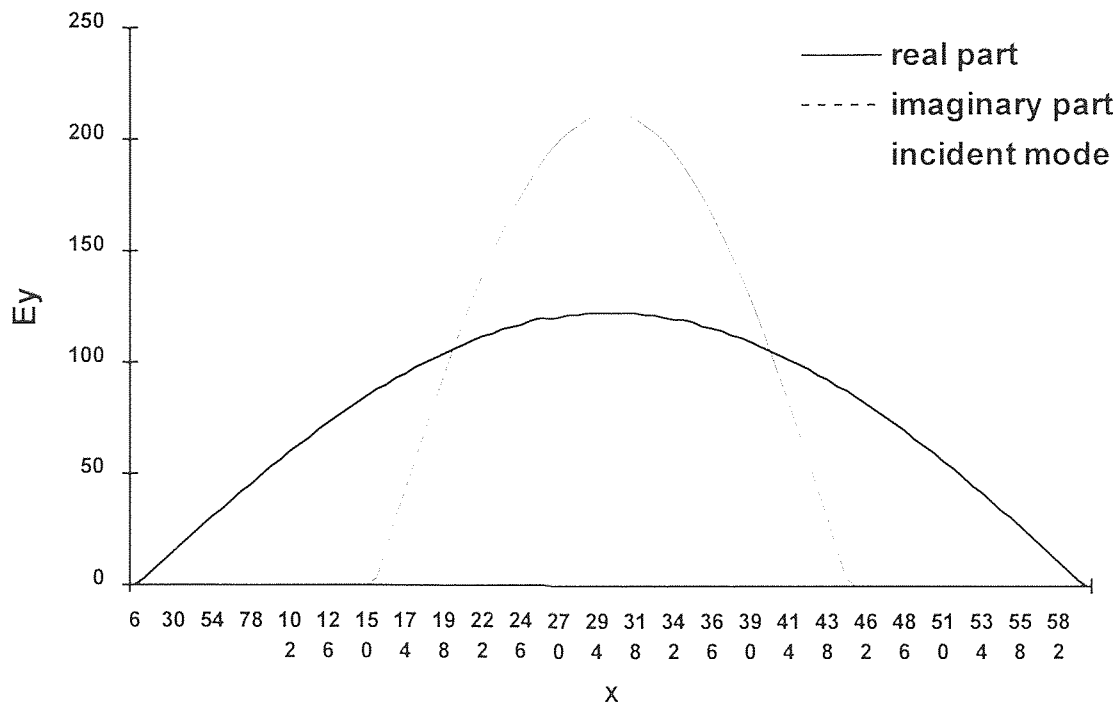
Figure 4.10   Electric field profiles at facet ten times b
away from the junction.
(a=1um, b=2um, λ=1.68um, 12 mode approximation)

To summarize the above analysis, a narrow stripe, so narrow that only propagating mode is the fundamental mode, is needed to confine the laser operating on single mode. That is why only fundamental modes are considered incident from there. The wide expanded stripe between the facet and junction need to be wide so that the light spot can be enlarged. But it can not be too wide. If the stripe is wide enough to hold another symmetric mode besides fundamental mode, multimode property will exist. It is no problem to make the stripe wide enough to hold the first antisymmetric mode besides a symmetric fundamental mode. That puts a limit for how wide the lateral expansion can be made. The conclusion is obviously general. The dielectric waveguides should also obey the rule.
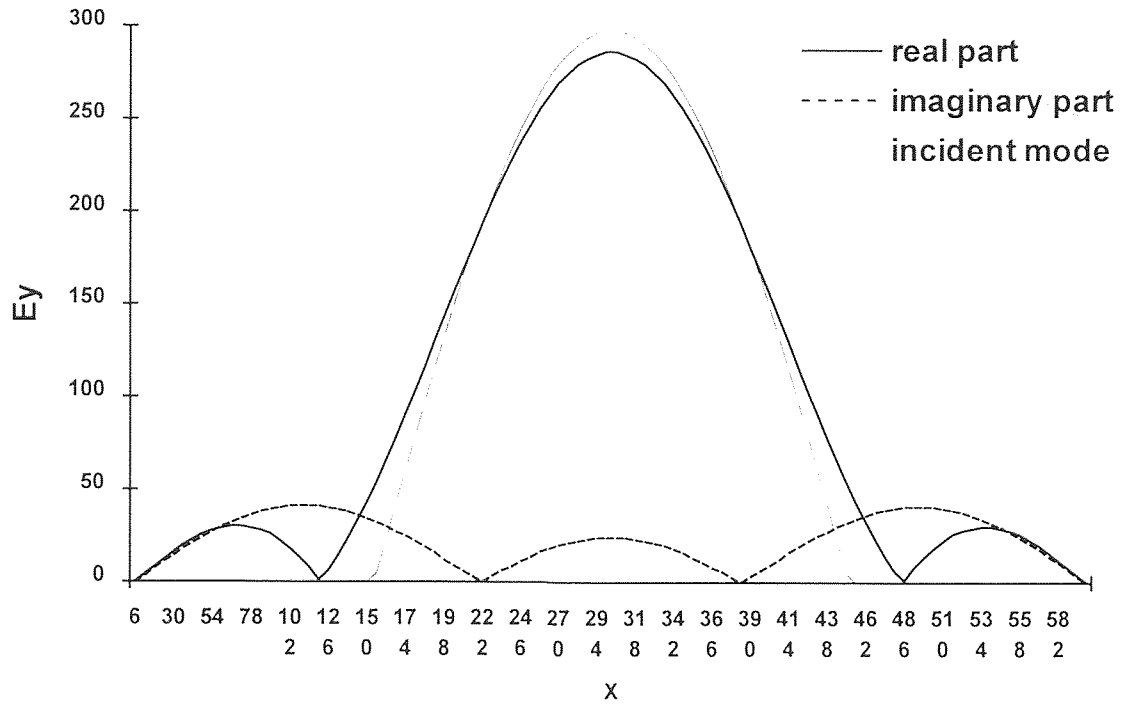
Figure 4.11 Electric field profiles at facet ten times b away from the junction.
(a=1um, b=2um, λ=0.88um, 12 mode approximation)
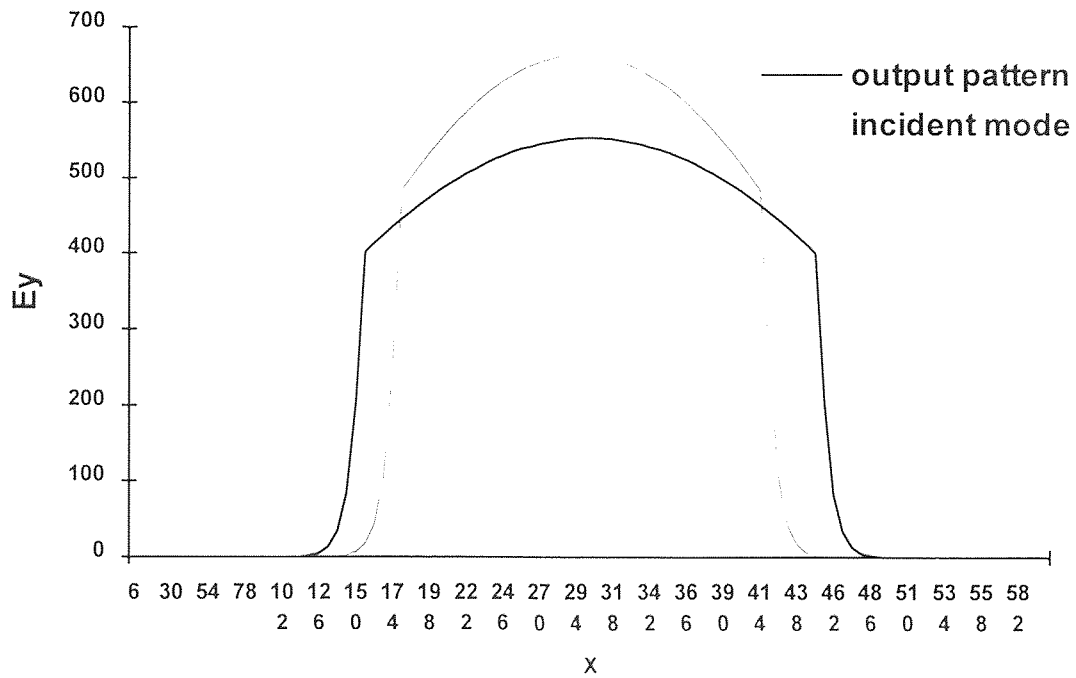
Figure 4.12 Electric field profiles at facet ten times b away from the junction.
(a=4um, b=5um, λ=0.88um, 5 mode approximation)

Up to now, the gain profile is not included. As presented in the next chapter, under certain gain profiles, the above requirement can be overcome. That brings a lot of possibilities for device design.

CHAPTER 5

CONCLUSION AND SUMMARY

The analysis of transverse discontinuity problems in waveguides using computer simulation, shows that the lateral expansion laser structure has its significance under certain circumstances. As studied in this thesis, not only the width of center part narrow stripe has a limit in keeping single mode operation, the expanded stripe ends also have a limit in width. The wide part can be wide enough to support the first antisymmetric mode, but not wide enough to hold another symmetric mode beside the fundamental mode.

The requirement is to make the higher order modes excited at the discontinuity junction evanescent modes. They decay as they propagate. So once the requirement is satisfied the higher order mode will not be able to reach the output facet.
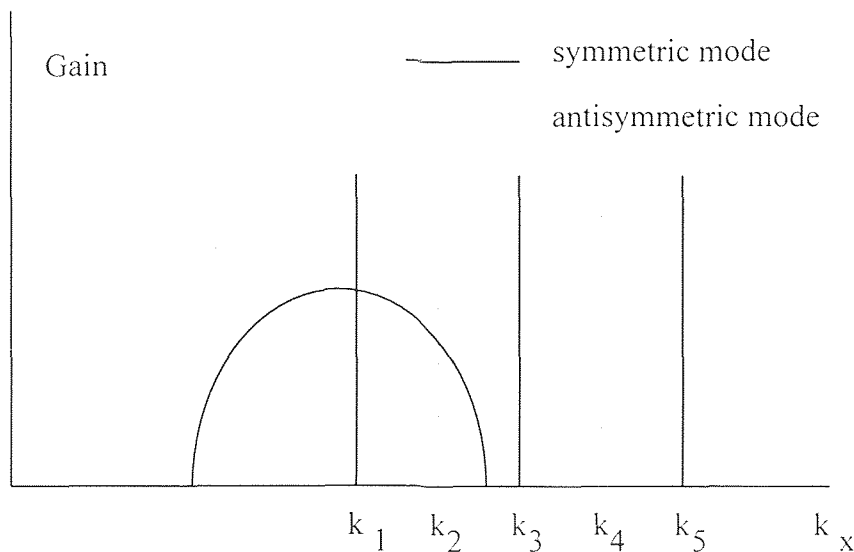
Figure 5.1 Gain profile

Of course the gain for each mode is not the same. If the gain profile is as shown in Figure 5.1, for which only the fundamental mode and the first antisymmetric mode has gain in the wide stripe, expansion stripes can be wider than we just discussed.

Using the gain profile shown in Figure 5.1, the structure discussed in chapter 4 emitting 0.88um light, can also achieve the performance of the lateral expansion structure as shown in Figure 5.2.
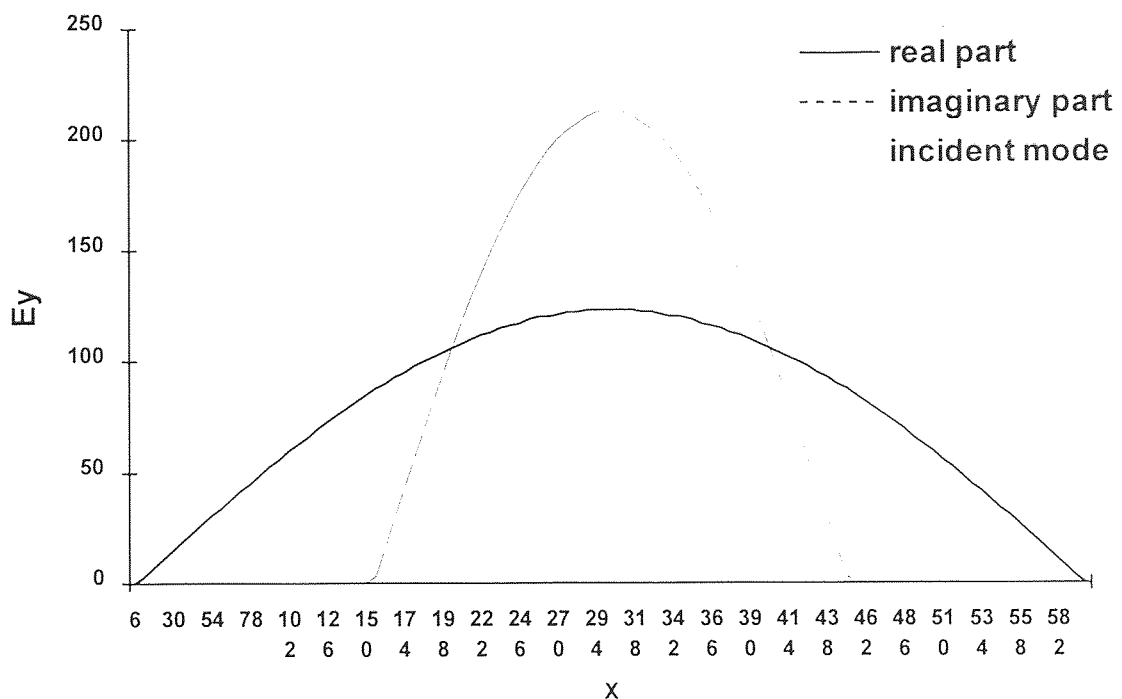
Figure 5.2 Electric field profiles at facet ten times b away
from the junction with gain profile as shown in Figure 5.1.
(a=1um, b=2um, λ=0.88um, 12 mode approximation)

If the gain profile can be as ideal as shown in Figure 5.3, there is no need for lateral expansion. Even wide stripe lasers can operate in a single mode. Of course, this gain profile is impractical for current technology.
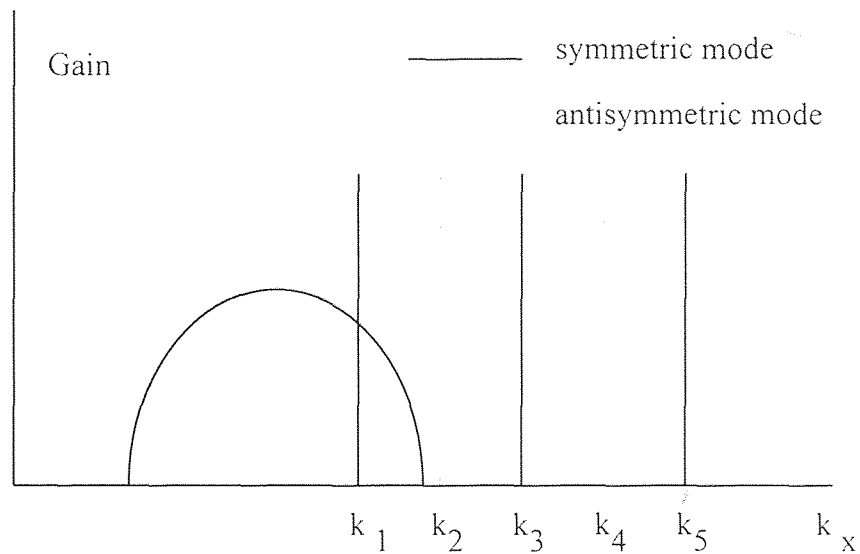
Figure 5.3 Ideal gain profile

Due to time problems, the radiation mode in dielectric waveguide is not included, which made the result for dielectric less accurate. The difficulty is in solving the eigenvalue equation of k_x, when k_x is complex variable. It can be a future work. Moreover, reflections happened within the cavity need to be included for more precise result.

SOURCE CODE FOR SOLUTION OF
METAL WAVEGUIDE DISCONTINUITY JUNCTION

```fortran
      program matrix
c****************************************************
c* Purpose: Calculate the matrix elements       *
c*          for mode matching at metal waveguide *
c*          junctions                     *
c****************************************************
c -- numerical integration by simpson's rule
c -- variables for integration
      integer out
      double precision sum,resum,imsum,upper,lower,tol
c
c -- variables for function calculation
      double precision a,b,u,u0,eps,eps0,lemda
      integer modl1,modr1,modl2,modr2,modmax
      complex pll(24,24),prr(24,24),plr(24,24)
c
      integer pau
      common /inout/ out
      common a,b
c
c ************************************************
      data pi/3.14159265/, tol/1.0d-5/
c ************************************************
c
      call datain(a,b,u,u0,eps,eps0,lemda,modmax)
c
      out=6
      write(out,*)' simpson''s rule integration'
c
      do 525 modl1=1,modmax
      do 313 modl2=1,modmax
      if (.not.(modl1.eq.modl2)) goto 900
c *********** real part of left **************
      call simps(lower, upper, tol, sum,
```

```
       *     a,b,lemda,u,eps,2*modl1-1,2*modl2-1,1)
       resum = sum
c ********** imaginary part of left **********
       call simps(lower, upper, tol, sum,
       *     a,b,lemda,u,eps,2*modl1-1,2*modl2-1,2)
       imsum = sum
c
       pll(modl1,modl2)=cmplx(resum,imsum)
       goto 901
c
900    pll(modl1,modl2)=cmplx(0.0,0.0)
901    continue
c      read(*,1) pau
1      format(i2)
313    write(out,104) modl1,modl2,pll(modl1,modl2)
104    format(/' mod1=,mod2=,rearea=,imarea=',2i3,2g14.7/)
525    continue
c
       do 2525 modr1=1,modmax
       do 2313 modr2=1,modmax
       if (.not.(modr1.eq.modr2)) goto 800
c *********** real part of right **************
       call simps(lower, upper, tol, sum,
       *     a,b,lemda,u,eps,(2*modr1-1),(2*modr2-1),3)
       resum = sum
c ********** imaginary part of right **********
       call simps(lower, upper, tol, sum,
       *     a,b,lemda,u,eps,(2*modr1-1),(2*modr2-1),4)
       imsum = sum
c
       prr(modr1,modr2)=cmplx(resum,imsum)
c
       goto 801
c
800    prr(modr1,modr2)=cmplx(0.0,0.0)
801    continue
c      read(*,1) pau
2313   write(out,104) modr1,modr2,prr(modr1,modr2)
2525   continue
c
       do 3525 modl1=1,modmax
```

```
      do 3313 modr2=1,modmax
c *********** real part of right **************
      call simps(lower, upper, tol, sum,
     *    a,b,lemda,u,eps,2*modl1-1,2*modr2-1,5)
      resum = sum
c ********** imaginary part of right ***********
      call simps(lower, upper, tol, sum,
     *    a,b,lemda,u,eps,2*modl1-1,2*modr2-1,6)
      imsum = sum
c
      plr(modl1,modr2)=cmplx(resum,imsum)
c
c     read(*,1) pau
      write(out,104) modl1,modr2,plr(modl1,modr2)
3313  continue
3525  continue
c
      call dataout(pll,prr,plr,modmax,a,b,lemda,eps,u)
c
      end
c
      subroutine simps(lower, upper, tol, sum,
     *        a,b,lemda,u,eps,mod1,mod2,choice)
c*************************************************
c*-- numerical integration by simpson's rule    *
c*-- with end correction                 *
c*************************************************
c -- variables for integration
      integer out, pieces, i, p2, choice
      double precision x, delta, lower, upper, sum, tol
      double precision endsum, oddsum, sum1, evsum, endcor
c -- variables for function calculation
      complex hxm0,hym0,exm0,eym0,beta
      complex dhxm0,dhym0,dexm0,deym0
      complex hxm0a,hym0a,exm0a,eym0a,betaa
      complex dhxm0a,dhym0a,dexm0a,deym0a
      double precision a,b,u,eps,lemda,width1,width2
      integer mod1,mod2
c
      common /inout/ out
c
```

```
      call wideup(width1,width2,lower,upper,choice,a,b)
c
      pieces = 2
      delta = (upper - lower) / pieces
      call ehxy(width1,lemda,u,eps,mod1,hxm0,hym0,exm0,eym0,
     *       lower+delta,beta)
      call ehxy(width2,lemda,u,eps,mod2,hxm0a,hym0a,exm0a,eym0a,
     *       lower+delta,betaa)
c
      if ((choice.eq.1).or.(choice.eq.3).or.(choice.eq.5))
     *   oddsum = real(-eym0 * conjg(hxm0a))
      if ((choice.eq.2).or.(choice.eq.4).or.(choice.eq.6))
     *   oddsum = aimag(-eym0 * conjg(hxm0a))
c
      evsum = 0.0
      call ehxy(width1,lemda,u,eps,mod1,hxm0,hym0,exm0,eym0,
     *       lower,beta)
      call ehxy(width2,lemda,u,eps,mod2,hxm0a,hym0a,exm0a,eym0a,
     *       lower,betaa)
c
      if ((choice.eq.1).or.(choice.eq.3).or.(choice.eq.5))
     *   endsum = real(-eym0 * conjg(hxm0a))
      if ((choice.eq.2).or.(choice.eq.4).or.(choice.eq.6))
     *   endsum = aimag(-eym0 * conjg(hxm0a))
c
      call ehxy(width1,lemda,u,eps,mod1,hxm0,hym0,exm0,eym0,
     *       upper,beta)
      call ehxy(width2,lemda,u,eps,mod2,hxm0a,hym0a,exm0a,eym0a,
     *       upper,betaa)
c
      if ((choice.eq.1).or.(choice.eq.3).or.(choice.eq.5))
     *   endsum = endsum+real(-eym0*conjg(hxm0a))
      if ((choice.eq.2).or.(choice.eq.4).or.(choice.eq.6))
     *   endsum = endsum+aimag(-eym0*conjg(hxm0a))
c
      call dehxy(width1,lemda,u,eps,mod1,dhxm0,dhym0,dexm0,
     *       deym0,lower,beta)
      call ehxy(width1,lemda,u,eps,mod1,hxm0,hym0,exm0,
     *       eym0,lower,beta)
      call dehxy(width2,lemda,u,eps,mod2,dhxm0a,dhym0a,dexm0a,
     *       deym0a,lower,betaa)
```

```fortran
      call ehxy(width2,lemda,u,eps,mod2,hxm0a,hym0a,exm0a,
     *        eym0a,lower,betaa)
c
      if ((choice.eq.1).or.(choice.eq.3).or.(choice.eq.5))
     *   endcor = real(-eym0*conjg(dhxm0a)
     *                     -deym0*conjg(hxm0a))
      if ((choice.eq.2).or.(choice.eq.4).or.(choice.eq.6))
     *   endcor = aimag(-eym0*conjg(dhxm0a)
     *                     -deym0*conjg(hxm0a))
c
      call dehxy(width1,lemda,u,eps,mod1,dhxm0,dhym0,dexm0,
     *        deym0,upper,beta)
      call ehxy(width1,lemda,u,eps,mod1,hxm0,hym0,exm0,
     *        eym0,upper,beta)
      call dehxy(width2,lemda,u,eps,mod2,dhxm0a,dhym0a,dexm0a,
     *        deym0a,upper,betaa)
      call ehxy(width2,lemda,u,eps,mod2,hxm0a,hym0a,exm0a,
     *        eym0a,upper,betaa)
c
      if ((choice.eq.1).or.(choice.eq.3).or.(choice.eq.5))
     *   endcor =endcor-real(-eym0*conjg(dhxm0a)
     *                     -deym0*conjg(hxm0a))
      if ((choice.eq.2).or.(choice.eq.4).or.(choice.eq.6))
     *   endcor =endcor-aimag(-eym0*conjg(dhxm0a)
     *                     -deym0*conjg(hxm0a))
c
      sum = (endsum + 4.0 * oddsum) * delta / 3.0
c
c      write(out, 101) pieces, sum
5     pieces = pieces * 2
      p2 = pieces /2
      sum1 = sum
      delta = (upper - lower) / pieces
      evsum = evsum + oddsum
      oddsum = 0.0
      do 10 i=1, p2
      x = lower + delta * (2.0 * i - 1.0)
      call ehxy(width1,lemda,u,eps,mod1,
     *        hxm0,hym0,exm0,eym0,x,beta)
      call ehxy(width2,lemda,u,eps,mod2,
     *        hxm0a,hym0a,exm0a,eym0a,x,betaa)
```

```
c
      if ((choice.eq.1).or.(choice.eq.3).or.(choice.eq.5))
     *    oddsum = oddsum + real(-eym0*conjg(hxm0a))
      if ((choice.eq.2).or.(choice.eq.4).or.(choice.eq.6))
     *    oddsum = oddsum + aimag(-eym0*conjg(hxm0a))
c
10    continue
      sum = (7.0 * endsum + 16.0 * oddsum + 14.0 * evsum
     *    + endcor * delta) * delta / 15.0
c     write(out, 101) pieces, sum
      if ( (abs(sum).le.1d-10).and.(abs(sum1).le.1d-10)
     *    .and.(i.gt.16000)) goto 313
      if (abs(sum - sum1) .gt. abs(tol * sum)) goto 5
      return
101   format(1x, i7, d14.7)
313   write(6,522)
522   format(' Integration i>16000 approximation')
      if (i.gt.16000) sum=0.0
      return
      end
c
      subroutine ehxy(width,lemda,u,eps,mod,
     *             hxm0,hym0,exm0,eym0,x,beta)
c********************************************************
c*   Generate transverse mode function         *
c********************************************************
      double precision x,width,lemda,u,eps,fm0,freq,w,v,a,b
      integer mod
      complex hxm0,hym0,exm0,eym0,beta
c
      common /inout/ out
      common a,b
c
c
      v = 1.0 / sqrt( u * eps )
      fm0 = mod * v / 2.0 / width
      freq = 2.99796d8 / lemda
      w = 2.0 * 3.1415926 * freq
      if ( freq .gt. fm0 ) beta = cmplx( ( 2.0 * 3.1415926
     *    / v * sqrt ( freq * freq - fm0 * fm0 ) ), 0.0 )
      if ( freq .le. fm0 ) beta = cmplx( 0.0, ( 2.0 * 3.1425926
```

```
   *    / v * sqrt ( fm0 * fm0 - freq * freq )) )
c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c                                                    c
c                                                    c
c                                                    c
c                    p=1 incident only                     c
c                                                    c
c                                                    c
c                                                    c

      if ((width.eq.a).and.(.not.(mod.eq.1))
   *    .and.(.not.(a.eq.b))) beta = -1.0 * beta


c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c
      hxm0 = (0.0,1.0) * beta / mod / 3.1415926
   *     * width * dsin ( dble(mod) * 3.1415926
   *     * (x-0.5*width) / width )
c
      hym0 = (0.0,0.0)
c
      exm0 = (0.0,0.0)
c
      eym0 = -w * u / beta * hxm0
      return
      end
c
      subroutine dehxy(width,lemda,u,eps,mod,
   *            dhxm0,dhym0,dexm0,deym0,x,beta)
c
c***********************************************
c*   Generate derivative of mode function     *
c***********************************************
      double precision x,width,lemda,u,eps,fm0,freq,w,v,a,b
      integer mod
      complex dhxm0,dhym0,dexm0,deym0,beta
c
      common /inout/ out
      common a,b
c
c
```

```
      v = 1.0 / sqrt( u * eps )
      fm0 = mod * v / 2.0 / width
      freq = 2.99796d8 / lemda
      w = 2.0 * 3.1415926 * freq
      if ( freq .gt. fm0 ) beta = cmplx( ( 2.0 * 3.1415926
     *    / v * sqrt ( freq * freq - fm0 * fm0 ) ), 0 )
      if ( freq .le. fm0 ) beta = cmplx( 0, ( 2.0 * 3.1425926
     *    / v * sqrt ( fm0 * fm0 - freq * freq )) )
c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c                                           c
c                                           c
c                                           c
c                 p=1 incident only             c
c                                           c
c                                           c
c                                           c

      if ((width.eq.a).and.(.not.(mod.eq.1))
     *    .and.(.not.(b.eq.a))) beta = -1.0 * beta


c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c
      dhxm0 = (0.0,1.0) * beta * dcos(
     *        dble(mod) * 3.1415926 * (x-0.5*width) / width )
c
      dhym0 = (0.0,0.0)
c
      dexm0 = (0.0,0.0)
c
      deym0 = -w * u / beta * dhxm0
      return
      end
c
      subroutine datain(a,b,u,u0,eps,eps0,lemda,modmax)
c**********************************************
c*  Input configurations of the structure     *
c**********************************************
c -- variables for function calculation
      double precision a,b,u,u0,eps,eps0,lemda
      integer modmax
c
```

```fortran
c     write(6,205)
205   format(' Input u:')
c     read(*,202) u
      u0 = 4.0 * 3.1415926d-7
      u = 1.0 * u0
c     write(6,208)
208   format(' Input eps:')
c     read(*,202) eps
      eps0 = 8.854d-12
      eps = 1.0 * eps0
c
      write(6,201)
201   format(' Input strip width at incidence side IN MICRONS:')
      read(*,202) a
      a = a * 1.0d-6
202   format(f10.0)
      write(6,203)
203   format(' Input strip width at output side IN MICRONS:')
      read(*,202) b
      b = b * 1.0d-6
      write(6,204)
204   format(' Input wave length in vacuum IN MICRONs')
      read(*,202) lemda
      lemda = lemda * 1.0d-6
c
      write(6,206)
206   format(' Input mode number modmax:')
      read(*,207) modmax
207   format(i2)
c
      write(6,205)
      write(6,202) u
      write(6,208)
      write(6,202) eps
      write(6,201)
      write(6,202) a*1.0d6
      write(6,203)
      write(6,202) b*1.0d6
      write(6,204)
      write(6,202) lemda*1.0d6
      write(6,206)
```

```
      write(6,207) modmax
      end
c
      subroutine wideup(width1,width2,lower,upper,choice,a,b)
c*************************************************
c*   Determine the lower and upper limits      *
c*     for the integration                 *
c*************************************************
      double precision width1,width2,lower,upper,a,b,abmin
      integer choice
c
      abmin = a / 2.0
      if (a.gt.b) abmin = b / 2.0
c
      if (choice.eq.1) width1=a
      if (choice.eq.1)  width2=a
      if (choice.eq.1)   lower=-a/2.0
      if (choice.eq.1)   upper= a/2.0
c
      if (choice.eq.2) width1=a
      if (choice.eq.2)  width2=a
      if (choice.eq.2)   lower=-a/2.0
      if (choice.eq.2)   upper= a/2.0
c
      if (choice.eq.3) width1=b
      if (choice.eq.3)  width2=b
      if (choice.eq.3)   lower=-b/2.0
      if (choice.eq.3)   upper= b/2.0
c
      if (choice.eq.4) width1=b
      if (choice.eq.4)  width2=b
      if (choice.eq.4)   lower=-b/2.0
      if (choice.eq.4)   upper= b/2.0
c
      if (choice.eq.5) width1=a
      if (choice.eq.5)  width2=b
      if (choice.eq.5)   lower=-abmin
      if (choice.eq.5)   upper= abmin
c
      if (choice.eq.6) width1=a
      if (choice.eq.6)  width2=b
```

```fortran
      if (choice.eq.6)    lower=-abmin
      if (choice.eq.6)    upper= abmin
c
      return
      end
c
      subroutine dataout(pll,prr,plr,modmax,a,b,lemda,eps,u)
c***************************************************
c*   Write the results of matrix element to     *
c*      a file: pmndata.                    *
c***************************************************
      integer modmax,i,j
      complex pll(24,24),prr(24,24),plr(24,24)
      double precision a,b,lemda,eps,u
c
      open(5,file = 'pmndata',status='new',form='formatted')
      write(5,101) modmax
101   format(1x,i3)
      write(5,103) a,b,lemda,eps,u
103   format(1x,5d14.7)
      do 10 i=1,modmax
       do 20 j=1,modmax
       write(5,102) pll(i,j),prr(i,j),plr(i,j)
20     continue
10    continue
102   format(1x,6e14.7)
      close(5)
      return
      end
      program solve
c***************************************************
c* Solve the simultaneous equations for       *
c*    the coefficients of each mode          *
c***************************************************
c -- fortran program to solve
c -- simultaneous equations by gauss elimination
c
      logical error
      integer length, maxr, out
      real a(48,48), y(48), coef(48)
c
```

```
      integer modmax
      complex pll(24,24),prr(24,24),plr(24,24)
      double precision al,br,lemda,eps,u
c
      common /inout/ out, maxr
c
      out = 6
c------
      maxr = 24
c------
      write(out, 101)
      call datain(pll,prr,plr,modmax,al,br,lemda,eps,u)
      call input(a, y, length,pll,prr,plr,modmax)
      if (length .lt. 2) stop
      call gauss(a, y, coef, length, error)
      if (.not.error) call output(a, y, coef, length)
101   format(' simultaneous solution by gauss elimination')
      end
c
      subroutine input(a, y, n,pll,prr,plr,modmax)
c -- get values for n and array a,y
c
      integer p,n, out, i, j, m, ri, modmax
      real a(48,48), y(48), rea(24,24), ima(24,24), rey(24), imy(24)
c
      complex pll(24,24),prr(24,24),plr(24,24)
c
      common /inout/ out, maxr
c
c********************
      p=1
c********************
5     write(out, 105)
      n = modmax * 2
      m = n
      if (n .gt. maxr) goto 5
      if (n .lt. 2) return
      do 99 ri = 1, 2
      if (ri .eq. 1) write(out,991)
      if (ri .eq. 2) write(out,992)
      do 20 i = 1, modmax
```

```
      write(out,101) i
      do 10 j = 1, modmax
       write(out,102) j
       if (ri .eq. 2) goto 2
        rea(i,j)=real(plr(j,i))
        rea(i+modmax,j+modmax)=real(conjg(plr(j,i)))
        rea(i,j+modmax)=real(prr(j,i))
        rea(i+modmax,j)=real(-1.0*conjg(pll(j,i)))
        write(out,444) i,j
444       format('now i,j= ',i3,i3)
        a(2*i-1,2*j-1)=rea(i,j)
        a(2*i,2*j)=rea(i,j)
        a(2*(i+modmax)-1,2*(j+modmax)-1)
     *          =rea((i+modmax),(j+modmax))
        a(2*(i+modmax),2*(j+modmax))
     *          =rea((i+modmax),(j+modmax))
        a(2*i-1,2*(j+modmax)-1)=rea(i,(j+modmax))
        a(2*i,2*(j+modmax))=rea(i,(j+modmax))
        a(2*(i+modmax)-1,2*j-1)=rea((i+modmax),j)
        a(2*(i+modmax),2*j)=rea((i+modmax),j)
        goto 10
2        ima(i,j)=aimag(plr(j,i))
        ima(i+modmax,j+modmax)=aimag(conjg(plr(j,i)))
        ima(i,j+modmax)=aimag(prr(j,i))
        ima(i+modmax,j)=aimag(-1.0*conjg(pll(j,i)))
        a(2*i-1,2*j)=-ima(i,j)
        a(2*i,2*j-1)=ima(i,j)
        a(2*(i+modmax)-1,2*(j+modmax))
     *          =-ima((i+modmax),(j+modmax))
        a(2*(i+modmax),2*(j+modmax)-1)
     *          =ima((i+modmax),(j+modmax))
        a(2*i-1,2*(j+modmax))=-ima(i,(j+modmax))
        a(2*i,2*(j+modmax)-1)=ima(i,(j+modmax))
        a(2*(i+modmax)-1,2*j)=-ima((i+modmax),j)
        a(2*(i+modmax),2*j-1)=ima((i+modmax),j)
10     continue
      write(out,104)
      if (ri .eq. 2) goto 22
       rey(i)=real(-plr(p,i))
       rey(i+modmax)=0.0
       if (i.eq.p) rey(i+modmax)=real(-conjg(pll(p,p)))
```

```fortran
      y(2*i-1)=rey(i)
      y(2*(i+modmax)-1)=rey(i+modmax)
      goto 20
22    imy(i)=aimag(-plr(p,i))
      imy(i+modmax)=0.0
      if (i.eq.p) imy(i+modmax)=aimag(-conjg(pll(p,p)))
      y(2*i)=imy(i)
      y(2*(i+modmax))=imy(i+modmax)
20    continue
99    continue
      do 111 i=1,n
111      write(out,333) (rea(i,j),j=1,n),rey(i)
      do 222 i=1,n
222      write(out,333) (ima(i,j),j=1,n),imy(i)
333      format(1p7e12.4)
      n = 2 * n
      return
991   format('----------Input real part----------')
992   format('----------Input imag part----------')
101   format(' equation ', i3/)
102   format('+',i4,': ' )
103   format(f10.0)
104   format('+ c: ' )
105   format(/'+how many equations? ' )
106   format(i2)
      end
c
      subroutine gauss(a, y, coef, ncol, error)
c************************************************
c -- simultaneous solution by gauss elimination *
c************************************************
      logical error
      integer ncol, i, j, out, n, n1, i1, k, l
      real a(48,48), y(48), coef(48)
      real b(48,48), w(48), big, ab, sum, t
      common /inout/ out, maxr
c
      error = .false.
      n = ncol
      do 20 i = 1 , n
      do 10 j = 1 , n
```

```fortran
        b(i,j) = a(i,j)
10    continue
      w(i) = y(i)
20    continue
      n1 = n - 1
      do 70 i = 1, n1
      big = abs(b(i,i))
      l = i
      i1 = i + 1
      do 30 j = i1, n
      ab = abs(b(j,i))
      if (ab .le. big) goto 30
       big = ab
       l = j
30    continue
      if (big .eq. 0.0) goto 70
      if (l .eq. i) goto 50
c -- interchange rows to put largest element on diagonal
      do 40 j=1, n
       call swap(b(l,j), b(i,j))
40    continue
      call swap(w(i),w(l))
50    continue
      do 60 j = i1, n
      t = b(j,i) / b(i,i)
      do 55 k = i1, n
       b(j,k) = b(j,k) - t * b(i,k)
55     continue
      w(j) = w(j) - t * w(i)
60    continue
70    continue
      if (b(n,n) .eq. 0.0) goto 99
      coef(n) = w(n) / b(n,n)
      i = n-1
c -- back substitution
80    sum = 0
      i1 = i + 1
      do 75 j = i1, n
      sum = sum + b(i,j) * coef(j)
75    continue
      coef(i) = (w(i)-sum) / b(i,i)
```

```
      i = i-1
      if (i .gt. 0) goto 80
      return
99    write(out,*)' error--matrix singular'
      return
      end
c
      subroutine swap(a,b)
c********************************************************
c -- interchange two values              *
c********************************************************
      real a, b, hold
c
      hold = a
      a = b
      b = hold
      return
      end
c
      subroutine output(a, y, coef, n)
c********************************************************
c -- Write the answers to a file: yidata.     *
c********************************************************
c
      integer n, out, i, j
      real a(48,48), y(48), coef(48),coabs2
      common /inout/ out, maxr
c
      open(5,file='yidata',status='new',form='formatted')
ct    write(5,*)'  Matrix elements:'
ct    write(5,*)'-----------------------------------------------'
      do 10 i = 1,n
      write(5, 101) (a(i,j), j =1, n), y(i)
10    continue
ct    write(5,102)
ct    write(5,*)'  Solution'
ct    write(5,*)' ----- only odd modes are count in case of p=1 -----'
ct    write(5,*)'-----------------------------------------------'
ct    write(5,*)'  real part   imag part   |a(i,p)|^2'
      do 900 i=1,n/2
ct        if (i.eq.1) write(5,*)' Left side:'
```

```
ct        if (i.eq.(n/4+1)) write(5,*)' Right side:'
          coabs2=coef(2*i-1)**2.0 + coef(2*i)**2.0
          write(5, 101) coef(2*i-1), coef(2*i), coabs2
900   continue
      close(5)
      return
101   format(1p3e14.6)
102   format(/)
      end
c
      subroutine datain(pll,prr,plr,modmax,a,b,lemda,eps,u)
c****************************************************
c* Read in the matrix elements            *
c****************************************************
      integer modmax,i,j
      complex pll(24,24),prr(24,24),plr(24,24)
      double precision a,b,lemda,eps,u
c
      open(4,file = 'pmndata',status='old',form='formatted')
      read(4,101) modmax
101   format(1x,i3)
      read(4,103) a,b,lemda,eps,u
103   format(1x,5d14.7)
      do 10 i=1,modmax
        do 20 j=1,modmax
        read(4,102) pll(i,j),prr(i,j),plr(i,j)
20      continue
10    continue
102   format(1x,6e14.7)
      close(5)
c
      write(6,101) modmax
      write(6,103) a,b,lemda,eps,u
      do 110 i=1,modmax
        do 120 j=1,modmax
        write(6,102) pll(i,j),prr(i,j),plr(i,j)
120     continue
110   continue
      end
      program an
c****************************************************
```

```
c*  Analysis the electric field profile      *
c*     at the metal waveguide junction         *
c*********************************************
      integer modmax,i,j,points,numx,numy
      complex pll(24,24),prr(24,24),plr(24,24),cmn(24)
      double precision max,max2,al,b,lemda,eps,u
      double precision imdiv,imet,rediv,rect
c
c -- variables for function calculation
      double precision widl, widr
      double precision wid,x
      integer modl,modr
      complex etl(300),etr(300),hxm0,hym0,exm0,eym0,beta
c
      integer n, out
      real a(48,48), y(48), coef(48),coabs2(24)
      common /inout/ out, maxr
      common al,b
c
      open(4,file = 'pmndata',status='old',form='formatted')
      read(4,101) modmax
101   format(1x,i3)
      read(4,103) al,b,lemda,eps,u
103   format(1x,5d14.7)
      do 10 i=1,modmax
        do 20 j=1,modmax
        read(4,102) pll(i,j),prr(i,j),plr(i,j)
20       continue
10    continue
102   format(1x,6e14.7)
      close(5)
c
      write(6,101) modmax
      write(6,103) al,b,lemda,eps,u
      do 110 i=1,modmax
        do 120 j=1,modmax
c       write(6,102) pll(i,j),prr(i,j),plr(i,j)
120      continue
110    continue
c
c
```

```
c
      n=4*modmax
      open(5,file='yidata',status='old',form='formatted')
      do 210 i = 1,n
       read(5, 201) (a(i,j), j =1, n), y(i)
210   continue
      do 900 i=1,n/2
        read(5, 201) coef(2*i-1), coef(2*i), coabs2(i)
900   cmn(i)=cmplx(coef(2*i-1), coef(2*i))
      close(5)
c
c     write(6,*)'  Matrix elements:'
c     write(6,*)'-----------------------------------------------'
      do 310 i = 1,n
c     write(6, 201) (a(i,j), j =1, n), y(i)
310   continue
c     write(6,1102)
      write(6,*)'  Solution'
      write(6,*)'-----------------------------------------------'
      write(6,*)' real part    imag part     |a(i,p)|^2'
        write(6,*) '  Left side wave guide:'
      do 1900 i=1,n/2
        if (i.eq.(n/4+1)) write(6,*) '  Right side wave guide:'
1900     write(6, 201) cmn(i),coabs2(i)
201   format(1p3e14.6)
1102  format(/)
c
c
      widl=al
      widr=b
c_____
c     write(*,1072)
c1072 format('  Input point number for analysis:')
c     read(*,1702) points
1702  format(i3)
      points=100
c_____
      max=0.0
      max2=0.0
      wid=al
      if (al.gt.b) wid=b
```

```
c     write(6,*) '  x= ,etl= ,etr= ,rediv= ,imdiv='
c
      do 313 i=0,(points-1)
      x=i*wid/(points-1)-wid/2.0
      call ehxy(al,lemda,u,eps,1,
     *          hxm0,hym0,exm0,eym0,x,beta)
      etl(i+1)=eym0*((1.0,0.0)+cmn(1))
      call ehxy(b,lemda,u,eps,1,
     *          hxm0,hym0,exm0,eym0,x,beta)
      etr(i+1)=eym0*cmn(1+modmax)
991   format(1x,1d14.7,4e12.5,2f7.2)
      do 525 modl=2,modmax
        call ehxy(al,lemda,u,eps,2*modl-1,
     *          hxm0,hym0,exm0,eym0,x,beta)
        etl(i+1)=etl(i+1)+eym0*cmn(modl)
        call ehxy(b,lemda,u,eps,2*modl-1,
     *          hxm0,hym0,exm0,eym0,x,beta)
        etr(i+1)=etr(i+1)+eym0*cmn(modl+modmax)
525   continue
c
      reet=real(etl(i+1))
      if (abs(real(etr(i+1))).gt.abs(real(etl(i+1))))
     *    reet=-real(etr(i+1))
      rediv=0
      if (abs(reet).gt.1e-12) rediv=real(etl(i+1)+etr(i+1))/reet
c
c
      imet=aimag(etl(i+1))
      if (abs(aimag(etr(i+1))).gt.abs(aimag(etl(i+1))))
     *    imet=-aimag(etr(i+1))
      imdiv=0
      if (abs(imet).gt.1e-12) imdiv=aimag(etl(i+1)+etr(i+1))/imet
c
c
      if ((i+1).eq.1) goto 1969
      if ((i+1).eq.points) goto 1969
      if (max.lt.abs(aimag(etl(i+1)))) max=abs(aimag(etl(i+1)))
      if (max2.lt.abs(aimag(etr(i+1)))) max2=abs(aimag(etr(i+1)))
      if (max3.lt.abs(real(etl(i+1)))) max3=abs(real(etl(i+1)))
      if (max4.lt.abs(real(etr(i+1)))) max4=abs(real(etr(i+1)))
1969  continue
```

```
c     write(6,991) x,etl(i+1),etr(i+1),rediv,imdiv
313   continue
c
      if (max.lt.max2) max=max2
      if (max3.lt.max4) max3=max4
      if (max.lt.1d-10) max=1.0
      if (max3.lt.1d-10) max3=1.0
      if (max.lt.max3) max=max3
      if (max3.lt.max) max3=max
c
      open(1,file='data1',status='new',form='formatted')
      open(5,file='data',status='new',form='formatted')
      open(4,file='datax',status='new',form='formatted')
      open(7,file='datarr',status='new',form='formatted')
      open(8,file='datari',status='new',form='formatted')
      open(9,file='datalr',status='new',form='formatted')
      open(10,file='datali',status='new',form='formatted')
      write(5,3333) points
3333  format(i4)
      do 4444 i=1,points
       numx = i * 600.0/real(points)
       numy = abs(aimag(etl(i)))*400.0/real(max)
       write(5,3333) numx
       write(4,3333) numx
       write(5,3333) numy
       write(10,3333) numy
4444  continue
      close(10)
      do 5555 i=1,points
       numx = i * 600.0/real(points)
       numy = abs(aimag(etr(i)))*400.0/real(max)
       write(5,3333) numx
       write(5,3333) numy
       write(8,3333) numy
5555  continue
      close(8)
      do 6666 i=1,points
       numx = i * 600.0/real(points)
       numy = abs(real(etl(i)))*400.0/real(max3)
       write(5,3333) numx
       write(5,3333) numy
```

```
         write(9,3333) numy
6666  continue
      close(9)
      do 7777 i=1,points
       numx = i * 600.0/real(points)
       numy = abs(real(etr(i)))*400.0/real(max3)
       write(5,3333) numx
       write(5,3333) numy
       write(7,3333) numy
7777  continue
      close(7)
      do 8888 i=1,points
       numx = (i-1) * 600.0/real(points)
       x=i*wid/(points-1)-wid/2.0
       call ehxy(al,lemda,u,eps,1,
     *           hxm0,hym0,exm0,eym0,x,beta)
       numy = cabs(eym0)*400.0/real(max)
       write(5,3333) numx
       write(5,3333) numy
       write(1,3333) numy
8888  continue
      close(1)
      close(5)
c
c
      end
c
c
      subroutine ehxy(width,lemda,u,eps,mod,
     *           hxm0,hym0,exm0,eym0,x,beta)
c*****************************************************
c  Generate transverse mode function          *
c*****************************************************
      double precision x,width,lemda,u,eps,fm0,freq,w,v,a,b
      integer mod
      complex hxm0,hym0,exm0,eym0,beta
c
      common a,b
c
c
      v = 1.0 / sqrt( u * eps )
```

```
      fm0 = mod * v / 2.0 / width
      freq = 2.99796d8 / lemda
      w = 2.0 * 3.1415926 * freq
      if ( freq .gt. fm0 ) beta = cmplx( ( 2.0 * 3.1415926
     *    / v * sqrt ( freq * freq - fm0 * fm0 ) ), 0.0 )
      if ( freq .le. fm0 ) beta = cmplx( 0.0, ( 2.0 * 3.1425926
     *    / v * sqrt ( fm0 * fm0 - freq * freq )) )
c
c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c                                          c
c                                          c
c                                          c
c              p=1 incident only                  c
c                                          c
c                                          c
c                                          c

      if ((width.eq.a).and.(.not.(mod.eq.1))
     *    .and.(.not.(a.eq.b))) beta = -1.0 * beta

c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c     write(6,*) 'I am here.'
      hxm0 = (0.0,1.0) * beta / mod / 3.1415926
     *    * width * dsin ( dble(mod) * 3.1415926
     *    * (x-0.5*width) / width )
c
      hym0 = (0.0,0.0)
c
      exm0 = (0.0,0.0)
c
      eym0 = -w * u / beta * hxm0
c     write(6,*) 'I passed ehxy.'
      return
      end
c
      program anend
c*********************************************************
c*  Analysis the electric field profile       *
c*    at the metal waveguide facet          *
c*********************************************************
c
```

```
      integer modmax,i,j,points,numx,numy
      complex pll(24,24),prr(24,24),plr(24,24),cmn(24)
      double precision max,max2,al,b,lemda,eps,u
      double precision imdiv,imet,rediv,reet
c
c -- variables for function calculation
      double precision widl, widr
      double precision wid,x,widmax
      integer modl,modr
      complex etl(300),etr(300),hxm0,hym0,exm0,eym0,beta
c
      integer n, out
      real a(48,48), y(48), coef(48),coabs2(24)
      common /inout/ out, maxr
      common al,b
c
      open(4,file = 'pmndata',status='old',form='formatted')
      read(4,101) modmax
101   format(1x,i3)
      read(4,103) al,b,lemda,eps,u
103   format(1x,5d14.7)
      do 10 i=1,modmax
       do 20 j=1,modmax
       read(4,102) pll(i,j),prr(i,j),plr(i,j)
20      continue
10    continue
102   format(1x,6e14.7)
      close(5)
c
c     write(6,101) modmax
c     write(6,103) al,b,lemda,eps,u
      do 110 i=1,modmax
       do 120 j=1,modmax
c       write(6,102) pll(i,j),prr(i,j),plr(i,j)
120      continue
110    continue
c
c
c
      n=4*modmax
      open(5,file='yidata',status='old',form='formatted')
```

```
      do 210 i = 1,n
        read(5, 201) (a(i,j), j =1, n), y(i)
210   continue
      do 900 i=1,n/2
        read(5, 201) coef(2*i-1), coef(2*i), coabs2(i)
900   cmn(i)=cmplx(coef(2*i-1), coef(2*i))
      close(5)
c
c     write(6,*)'  Matrix elements:'
c     write(6,*)'------------------------------------------------'
      do 310 i = 1,n
c       write(6, 201) (a(i,j), j =1, n), y(i)
310   continue
c     write(6,1102)
      write(6,*)'  Solution'
      write(6,*)'-----------------------------------------------'
      write(6,*)' real part    imag part    |a(i,p)|^2'
        write(6,*) '  Left side wave guide:'
      do 1900 i=1,n/2
        if (i.eq.(n/4+1)) write(6,*) '  Right side wave guide:'
1900    write(6, 201) cmn(i),coabs2(i)
201   format(1p3e14.6)
1102  format(/)
c
c
      widl=al
      widr=b
c_____
c     write(*,1072)
c1072  format('  Input point number for analysis:')
c     read(*,1702) points
1702  format(i3)
      points=100
c_____
      max=0.0
      max2=0.0
      wid=al
      widmax=b
      if (al.gt.b) wid=b
      if (al.gt.b) widmax=al
c     write(6,*) '  x= ,etl= ,etr= ,rediv= ,imdiv='
```

```
c
      do 313 i=0,(points-1)
      x=i*widmax/(points-1)-widmax/2.0
      call ehxy(al,lemda,u,eps,1,
     *          hxm0,hym0,exm0,eym0,x,beta)
      etl(i+1)=eym0*((1.0,0.0)+cmn(1))
        if ((al.lt.widmax).and.(x.gt.(wid/2.0))) etl(i+1)=0
        if ((al.lt.widmax).and.(x.lt.(-wid/2.0))) etl(i+1)=0
      call ehxy(b,lemda,u,eps,1,
     *          hxm0,hym0,exm0,eym0,x,beta)
      etr(i+1)=eym0*cmn(1+modmax)*sqrt
     * (cexp((0,1)*beta*10.0*b)*conjg(cexp((0,1)*beta*10.0*b)))
        if ((b.lt.widmax).and.(x.gt.(wid/2.0))) etr(i+1)=0
        if ((b.lt.widmax).and.(x.lt.(-wid/2.0))) etr(i+1)=0
991   format(1x,1d14.7,4e12.5,2f7.2)
      do 525 modl=2,modmax
       call ehxy(al,lemda,u,eps,2*modl-1,
     *          hxm0,hym0,exm0,eym0,x,beta)
       etl(i+1)=etl(i+1)+eym0*cmn(modl)
       if ((al.lt.widmax).and.(x.gt.(wid/2.0))) etl(i+1)=0
       if ((al.lt.widmax).and.(x.lt.(-wid/2.0))) etl(i+1)=0
       call ehxy(b,lemda,u,eps,2*modl-1,
     *          hxm0,hym0,exm0,eym0,x,beta)
       etr(i+1)=etr(i+1)+eym0*cmn(modl+modmax)*sqrt
     * (cexp((0,1)*beta*10.0*b)*conjg(cexp((0,1)*beta*10.0*b)))
       if ((b.lt.widmax).and.(x.gt.(wid/2.0))) etr(i+1)=0
       if ((b.lt.widmax).and.(x.lt.(-wid/2.0))) etr(i+1)=0
525   continue
c
      reet=real(etl(i+1))
      if (abs(real(etr(i+1))).gt.abs(real(etl(i+1))))
     *   reet=-real(etr(i+1))
      rediv=0
      if (abs(reet).gt.1e-12) rediv=real(etl(i+1)+etr(i+1))/reet
c
c
      imet=aimag(etl(i+1))
      if (abs(aimag(etr(i+1))).gt.abs(aimag(etl(i+1))))
     *   imet=-aimag(etr(i+1))
      imdiv=0
      if (abs(imet).gt.1e-12) imdiv=aimag(etl(i+1)+etr(i+1))/imet
```

```fortran
c
c
      if ((i+1).eq.1) goto 1969
      if ((i+1).eq.points) goto 1969
      if (max.lt.abs(aimag(etl(i+1)))) max=abs(aimag(etl(i+1)))
      if (max2.lt.abs(aimag(etr(i+1)))) max2=abs(aimag(etr(i+1)))
      if (max3.lt.abs(real(etl(i+1)))) max3=abs(real(etl(i+1)))
      if (max4.lt.abs(real(etr(i+1)))) max4=abs(real(etr(i+1)))
1969  continue
c     write(6,991) x,etl(i+1),etr(i+1),rediv,imdiv
313   continue
c
      if (max.lt.max2) max=max2
      if (max3.lt.max4) max3=max4
      if (max.lt.1d-10) max=1.0
      if (max3.lt.1d-10) max3=1.0
      if (max.lt.max3) max=max3
      if (max3.lt.max) max3=max
c
      open(1,file='data1',status='new',form='formatted')
      open(5,file='data',status='new',form='formatted')
      open(4,file='datax',status='new',form='formatted')
      open(7,file='datarr',status='new',form='formatted')
      open(8,file='datari',status='new',form='formatted')
      open(9,file='datalr',status='new',form='formatted')
      open(10,file='datali',status='new',form='formatted')
      write(5,3333) points
3333  format(i4)
      do 4444 i=1,points
        numx = i * 600.0/real(points)
        numy = abs(aimag(etl(i)))*400.0/real(max)
        write(5,3333) numx
        write(4,3333) numx
        write(5,3333) numy
        write(10,3333) numy
4444  continue
      close(10)
      do 5555 i=1,points
        numx = i * 600.0/real(points)
        numy = abs(aimag(etr(i)))*400.0/real(max)
        write(5,3333) numx
```

```
      write(5,3333) numy
      write(8,3333) numy
5555  continue
      close(8)
      do 6666 i=1,points
       numx = i * 600.0/real(points)
       numy = abs(real(etl(i)))*400.0/real(max3)
       write(5,3333) numx
       write(5,3333) numy
       write(9,3333) numy
6666  continue
      close(9)
      do 7777 i=1,points
       numx = i * 600.0/real(points)
       numy = abs(real(etr(i)))*400.0/real(max3)
       write(5,3333) numx
       write(5,3333) numy
       write(7,3333) numy
7777  continue
      close(7)
      do 8888 i=1,points
       numx = (i-1) * 600.0/real(points)
       x=(i-1)*widmax/(points-1)-widmax/2.0
       call ehxy(al,lemda,u,eps,1,
     *           hxm0,hym0,exm0,eym0,x,beta)
       numy = cabs(eym0)*400.0/real(max)
       if ((al.lt.widmax).and.(x.gt.(wid/2.0))) numy =0
       if ((al.lt.widmax).and.(x.lt.(-wid/2.0))) numy =0
       write(5,3333) numx
       write(5,3333) numy
       write(1,3333) numy
8888  continue
      close(1)
      close(5)
c
c
      end
c
c
      subroutine ehxy(width,lemda,u,eps,mod,
     *           hxm0,hym0,exm0,eym0,x,beta)
```

```
c***********************************************
c*  Generate transverse mode functions        *
c***********************************************
c
      double precision x,width,lemda,u,eps,fm0,freq,w,v,a,b
      integer mod
      complex hxm0,hym0,exm0,eym0,beta
c
      common a,b
c
c
      v = 1.0 / sqrt( u * eps )
      fm0 = mod * v / 2.0 / width
      freq = 2.99796d8 / lemda
      w = 2.0 * 3.1415926 * freq
      if ( freq .gt. fm0 ) beta = cmplx( ( 2.0 * 3.1415926
     *    / v * sqrt ( freq * freq - fm0 * fm0 ) ), 0.0 )
      if ( freq .le. fm0 ) beta = cmplx( 0.0, ( 2.0 * 3.1425926
     *    / v * sqrt ( fm0 * fm0 - freq * freq )) )
c
c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c                                        c
c                                        c
c                                        c
c              p=1 incident only              c
c                                        c
c                                        c
c                                        c

      if ((width.eq.a).and.(.not.(mod.eq.1))
     *    .and.(.not.(a.eq.b))) beta = -1.0 * beta

c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c    write(6,*) 'I am here.'
      hxm0 = (0.0,1.0) * beta / mod / 3.1415926
     *    * width * dsin ( dble(mod) * 3.1415926
     *    * (x-0.5*width) / width )
c
      hym0 = (0.0,0.0)
c
      exm0 = (0.0,0.0)
```

```
c
      eym0 = -w * u / beta * hxm0
c     write(6,*) 'I passed ehxy.'
      return
      end
c
      program aneach
c*************************************************
c*  Aalysis the first five modes' components of  *
c*    electric field profile                     *
c*    at the metal waveguide junction          *
c*************************************************
c
      integer modmax,i,j,points,numx,numy
      complex pll(24,24),prr(24,24),plr(24,24),cmn(24)
      double precision max,max2,al,b,lemda,eps,u
      double precision imdiv,imet,rediv,reet
c
c -- variables for function calculation
      double precision widl, widr
      double precision wid,x
      integer modl,modr
      complex etl(300),etr(300),hxm0,hym0,exm0,eym0,beta
c
      integer n, out
      real a(48,48), y(48), coef(48),coabs2(24)
      common /inout/ out, maxr
      common al,b
c
      open(4,file = 'pmndata',status='old',form='formatted')
      read(4,101) modmax
101   format(1x,i3)
      read(4,103) al,b,lemda,eps,u
103   format(1x,5d14.7)
      do 10 i=1,modmax
        do 20 j=1,modmax
        read(4,102) pll(i,j),prr(i,j),plr(i,j)
20      continue
10    continue
102   format(1x,6e14.7)
      close(5)
```

```
c
      write(6,101) modmax
      write(6,103) al,b,lemda,eps,u
      do 110 i=1,modmax
        do 120 j=1,modmax
c       write(6,102) pll(i,j),prr(i,j),plr(i,j)
120     continue
110   continue
c
c
c
      n=4*modmax
      open(5,file='yidata',status='old',form='formatted')
      do 210 i = 1,n
        read(5, 201) (a(i,j), j =1, n), y(i)
210   continue
      do 900 i=1,n/2
        read(5, 201) coef(2*i-1), coef(2*i), coabs2(i)
900   cmn(i)=cmplx(coef(2*i-1), coef(2*i))
      close(5)
c
c     write(6,*)'  Matrix elements:'
c     write(6,*)'----------------------------------------------------'
      do 310 i = 1,n
c       write(6, 201) (a(i,j), j =1, n), y(i)
310   continue
c     write(6,1102)
      write(6,*)'  Solution'
      write(6,*)'----------------------------------------------------'
      write(6,*)' real part    imag part    |a(i,p)|^2'
        write(6,*) '  Left side wave guide:'
      do 1900 i=1,n/2
        if (i.eq.(n/4+1)) write(6,*) '  Right side wave guide:'
1900    write(6, 201) cmn(i),coabs2(i)
201   format(1p3e14.6)
1102  format(/)
c
c
      widl=al
      widr=b
c_____
```

```
c     write(*,1072)
c1072  format('  Input point number for analysis:')
c     read(*,1702) points
1702  format(i3)
      points=100
c_____
      max=0.0
      max2=0.0
      wid=al
      if (al.gt.b) wid=b
c     write(6,*) '  x= ,etl= ,etr= ,rediv= ,imdiv='
c
      do 313 i=0,(points-1)
      x=i*wid/(points-1)-wid/2.0
      call ehxy(al,lemda,u,eps,1,
     *          hxm0,hym0,exm0,eym0,x,beta)
      etl(i+1)=eym0*((1.0,0.0)+cmn(1))
      call ehxy(b,lemda,u,eps,1,
     *          hxm0,hym0,exm0,eym0,x,beta)
      etr(i+1)=eym0*cmn(1+modmax)
991   format(1x,1d14.7,4e12.5,2f7.2)
      do 525 modl=2,modmax
       call ehxy(al,lemda,u,eps,2*modl-1,
     *          hxm0,hym0,exm0,eym0,x,beta)
       etl(i+1)=etl(i+1)+eym0*cmn(modl)
       call ehxy(b,lemda,u,eps,2*modl-1,
     *          hxm0,hym0,exm0,eym0,x,beta)
       etr(i+1)=etr(i+1)+eym0*cmn(modl+modmax)
525   continue
c
      reet=real(etl(i+1))
      if (abs(real(etr(i+1))).gt.abs(real(etl(i+1))))
     *    reet=-real(etr(i+1))
      rediv=0
      if (abs(reet).gt.1e-12) rediv=real(etl(i+1)+etr(i+1))/reet
c
c
      imet=aimag(etl(i+1))
      if (abs(aimag(etr(i+1))).gt.abs(aimag(etl(i+1))))
     *    imet=-aimag(etr(i+1))
      imdiv=0
```

```
      if (abs(imet).gt.1e-12) imdiv=aimag(etl(i+1)+etr(i+1))/imet
c
c
      if ((i+1).eq.1) goto 1969
      if ((i+1).eq.points) goto 1969
      if (max.lt.abs(aimag(etl(i+1)))) max=abs(aimag(etl(i+1)))
      if (max2.lt.abs(aimag(etr(i+1)))) max2=abs(aimag(etr(i+1)))
      if (max3.lt.abs(real(etl(i+1)))) max3=abs(real(etl(i+1)))
      if (max4.lt.abs(real(etr(i+1)))) max4=abs(real(etr(i+1)))
1969  continue
c     write(6,991) x,etl(i+1),etr(i+1),rediv,imdiv
313   continue
c
      if (max.lt.max2) max=max2
      if (max3.lt.max4) max3=max4
      if (max.lt.1d-10) max=1.0
      if (max3.lt.1d-10) max3=1.0
      if (max.lt.max3) max=max3
      if (max3.lt.max) max3=max
c
      open(1,file='data1',status='new',form='formatted')
      open(2,file='data2',status='new',form='formatted')
      open(3,file='data3',status='new',form='formatted')
      open(4,file='data4',status='new',form='formatted')
      open(5,file='data',status='new',form='formatted')
      open(8,file='datax',status='new',form='formatted')
      open(7,file='data5',status='new',form='formatted')
      write(5,3333) points
3333  format(i4)
      do 4444 i=1,points
       numx = i * 600.0/real(points)
       x=i*b/(points-1)-b/2.0
       call ehxy(b,lemda,u,eps,9,
     *           hxm0,hym0,exm0,eym0,x,beta)
       numy = cabs(cmn(5+modmax)*eym0)*400.0/real(max)
       write(5,3333) numx
       write(5,3333) numy
       write(8,3333) numx
       write(7,3333) numy
4444  continue
      do 5555 i=1,points
```

```
      numx = i * 600.0/real(points)
      x=i*b/(points-1)-b/2.0
      call ehxy(b,lemda,u,eps,7,
   *           hxm0,hym0,exm0,eym0,x,beta)
      numy = cabs(cmn(4+modmax)*eym0)*400.0/real(max)
      write(5,3333) numx
      write(5,3333) numy
      write(4,3333) numy
5555  continue
      do 6666 i=1,points
      numx = i * 600.0/real(points)
      x=i*b/(points-1)-b/2.0
      call ehxy(b,lemda,u,eps,5,
   *           hxm0,hym0,exm0,eym0,x,beta)
      numy = cabs(cmn(3+modmax)*eym0)*400.0/real(max)
      write(5,3333) numx
      write(5,3333) numy
      write(3,3333) numy
6666  continue
      do 7777 i=1,points
      numx = i * 600.0/real(points)
      x=i*b/(points-1)-b/2.0
      call ehxy(b,lemda,u,eps,3,
   *           hxm0,hym0,exm0,eym0,x,beta)
      numy = cabs(cmn(2+modmax)*eym0)*400.0/real(max)
      write(5,3333) numx
      write(5,3333) numy
      write(2,3333) numy
7777  continue
      do 8888 i=1,points
      numx = i * 600.0/real(points)
      x=i*b/(points-1)-b/2.0
      call ehxy(b,lemda,u,eps,1,
   *           hxm0,hym0,exm0,eym0,x,beta)
      numy = cabs(cmn(1+modmax)*eym0)*400.0/real(max)
      write(5,3333) numx
      write(5,3333) numy
      write(1,3333) numy
8888  continue
      close(1)
      close(2)
```

```
      close(3)
      close(4)
      close(5)
      close(7)
      close(8)
c
c
      end
c
c
      subroutine ehxy(width,lemda,u,eps,mod,
     *            hxm0,hym0,exm0,eym0,x,beta)
c
c**************************************************
c*  Generate the transverse mode functions      *
c**************************************************
c
      double precision x,width,lemda,u,eps,fm0,freq,w,v,a,b
      integer mod
      complex hxm0,hym0,exm0,eym0,beta
c
      common a,b
c
c
      v = 1.0 / sqrt( u * eps )
      fm0 = mod * v / 2.0 / width
      freq = 2.99796d8 / lemda
      w = 2.0 * 3.1415926 * freq
      if ( freq .gt. fm0 ) beta = cmplx( ( 2.0 * 3.1415926
     *   / v * sqrt ( freq * freq - fm0 * fm0 ) ), 0.0 )
      if ( freq .le. fm0 ) beta = cmplx( 0.0, ( 2.0 * 3.1425926
     *   / v * sqrt ( fm0 * fm0 - freq * freq )) )
c
c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c                                          c
c                                          c
c                                          c
c             p=1 incident only               c
c                                          c
c                                          c
c                                          c
```

```fortran
      if ((width.eq.a).and.(.not.(mod.eq.1))
     *    .and.(.not.(a.eq.b))) beta = -1.0 * beta

c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c     write(6,*) 'I am here.'
      hxm0 = (0.0,1.0) * beta / mod / 3.1415926
     *    * width * dsin ( dble(mod) * 3.1415926
     *    * (x-0.5*width) / width )
c
      hym0 = (0.0,0.0)
c
      exm0 = (0.0,0.0)
c
      eym0 = -w * u / beta * hxm0
c     write(6,*) 'I passed ehxy.'
      return
      end
c
```

```c
c**************************************************
c*          Purpose: Plot analysis results on the screen      *
c*          (program is writen in C language)          *
c**************************************************
#include <stdio.h>
#include <math.h>
#include <graphics.h>

int driver,mode;

main()
{
int i,points,numx[311],numy[311];
char *u;
FILE *ifl;

driver=DETECT;
initgraph(&driver,&mode,"");
cleardevice();
```

```
ifl=fopen("data","r");

  setcolor(EGA_LIGHTBLUE);
  fgets(u,10,ifl);
  points=atoi(u);
  printf("%d\n",points);
  fgets(u,10,ifl);
  numx[1]=atoi(u);
  fgets(u,10,ifl);
  numy[1]=atoi(u);
for (i=2;i<(points+1);i++)
{
  fgets(u,10,ifl);
  numx[i]=atoi(u);
  fgets(u,10,ifl);
  numy[i]=atoi(u);
  line(numx[i-1],450-numy[i-1]+1,numx[i],450-numy[i]+1);
};

  setcolor(EGA_GREEN);
  fgets(u,10,ifl);
  numx[1]=atoi(u);
  fgets(u,10,ifl);
  numy[1]=atoi(u);
for (i=2;i<(points+1);i++)
{
  fgets(u,10,ifl);
  numx[i]=atoi(u);
  fgets(u,10,ifl);
  numy[i]=atoi(u);
  line(numx[i-1],450-numy[i-1],numx[i],450-numy[i]);
};

  setcolor(EGA_LIGHTRED);
  fgets(u,10,ifl);
  numx[1]=atoi(u);
  fgets(u,10,ifl);
  numy[1]=atoi(u);
for (i=2;i<(points+1);i++)
{
```

```
  fgets(u,10,ifl);
  numx[i]=atoi(u);
  fgets(u,10,ifl);
  numy[i]=atoi(u);
  line(numx[i-1],450-numy[i-1]+1,numx[i],450-numy[i]+1);
};

  setcolor(EGA_RED);
  fgets(u,10,ifl);
  numx[1]=atoi(u);
  fgets(u,10,ifl);
  numy[1]=atoi(u);
 for (i=2;i<(points+1);i++)
 {
  fgets(u,10,ifl);
  numx[i]=atoi(u);
  fgets(u,10,ifl);
  numy[i]=atoi(u);
  line(numx[i-1],450-numy[i-1],numx[i],450-numy[i]);
};

  setcolor(EGA_WHITE);
  fgets(u,10,ifl);
  numx[1]=atoi(u);
  fgets(u,10,ifl);
  numy[1]=atoi(u);
 for (i=2;i<(points+1);i++)
 {
  fgets(u,10,ifl);
  numx[i]=atoi(u);
  fgets(u,10,ifl);
  numy[i]=atoi(u);
  line(numx[i-1],450-numy[i-1],numx[i],450-numy[i]);
};

fclose(ifl);
}
```

APPENDIX B

SOURCE CODE FOR SOLUTION
DIELECTRIC WAVEGUIDE DISCONTINUITY JUNCTION

Except for the transverse mode function, as presented below, other parts of the program for dielectric waveguide junction are the same as the program for metal waveguide junction.

```
      subroutine ehxy(width,lemda,u,eps,mod,
     *            hxm0,hym0,exm0,eym0,x,beta)
c*****************************************************
c*  Generate the transverse mode functions      *
c*****************************************************
c
      double precision x,width,lemda,u,eps,fm0,freq,w,v,a,b
      double precision Am0,Bm0,kx,epsclad,epscore
      integer mod,modtreat
      complex hxm0,hym0,exm0,eym0,beta,alphax
c
      common /inout/ out
      common a,b
c
      epsclad = eps * 11.4921
      epscore = eps * 13.0321
      freq = 2.99796d8 / lemda
      w = 2.0 * 3.1415926 * freq
      modtreat=sqrt(w*w*u*(width/2.0)*(width/2.0)
     *    *(epscore-epsclad))/3.1415926-0.5+0.5+1.0
c
      kx=0.0
      Am0=1.0
      if (mod.gt.(2*modtreat-1)) Am0 = 0.0
      if (real(mod/2).eq.(real(mod)/2.0)) Am0=0.0
c     if (mod.gt.(2*modtreat-1)) goto 7029
      call root(epscore,epsclad,u,width,lemda,kx,mod)
c
      beta = csqrt(w*w*u*epscore - kx*kx)
      alphax = csqrt(beta*beta - w*w*u*epsclad)
```

```
c
c
c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c                                              c
c                                              c
c                                              c
c                 p=1 incident only              c
c                                              c
c                                              c
c                                              c

      if ((width.eq.a).and.(.not.(mod.eq.1))
     *    .and.(.not.(b.eq.a))) beta = -1.0 * beta


c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c c
c
      if (x.lt.(-width/2.0)) goto 1
      if (x.gt.(width/2.0)) goto 3
      goto 2
c
1     eym0 =Am0*cexp(alphax*x)/cexp(alphax*(-width))/2.0)
      exm0 = (0.0,0.0)
      hym0 = (0.0,0.0)
      hxm0 = -beta / w / u * eym0
      goto 9
c
2     eym0 = Am0 * cos(kx*x) / cos(kx*width/2.0)
      exm0 = (0.0,0.0)
      hym0 = (0.0,0.0)
      hxm0 = -beta / w / u * eym0
      goto 9
c
3     eym0 = Am0 * cexp(-alphax*x)/cexp(alphax*(-width)/2.0)
      exm0 = (0.0,0.0)
      hym0 = (0.0,0.0)
      hxm0 = -beta / w / u * eym0
9     return
c
7029  eym0= (0.0,0.0)
      exm0 = (0.0,0.0)
      hxm0 = (0.0,0.0)
```

```fortran
      hym0 = (0.0,0.0)
      return
c
      end
c
      subroutine root(epscore,epsclad,u,width,lemda,kx,mod)
c**************************************************
c*  Solving eigenvalue equation of propagation constant  *
c**************************************************
c
      double precision epscore,epsclad,u,width,lemda
      double precision left,right,div,kx,x,y,xmax
      integer mod
c
      div=1d-5
      w=2.0*3.1415926*3.0e8/lemda
      xmax=2*sqrt(w*w*u*(width/2.0)*(width/2.0)*(epscore-epsclad))
      left= (real((mod-3)/2.0)+0.5)*3.1415926+1d-8
      right=(real((mod-1)/2.0)+0.5)*3.1415926-1d-8
      if (mod.eq.1) left=0.0+1d-8
      if (mod.eq.1) right=0.5*3.1415926-1d-8
      if (xmax.lt.right) right=xmax
10    x=(left+right)/2.0
      y=tan(x)-sqrt(xmax*xmax/x/x-1)
      if (y.eq.0.0) kx=x/width
      if (y.eq.0.0) return
      if (y.lt.0.0) left=x
      if (y.gt.0.0) right=x
      if ((right-left).gt.div) goto 10
      kx=x/width
      return
      end
```

BIBLIOGRAPHY

Adams, Lee. 1988. *High Performance Graphics in C Animation and Simulation.* Winderest.

Anderson, D. B., J. T. Boyd, and J. D. Mullen. 1970. "Dielectric Waveguide Phase Matching of Infrared Parametric Interactions." *Proceeding of the Polytechnic Institute of Brooklyn Symposium on Submillimeter Waves*: 191-209.

Argence, E., and T. Kahan. 1968. *Theory of Waveguides and Cavity Resonators.* Hart Pub. Co.

Ashcroft, J., R. H. Eldridge, R. W. Paulson and G. A. Wilson. 1981. *Programming with Fortran 77.* Collins; Distributed in U.S.A. by Sheridan House.

Clarricoats, P.J.B., and A. B. Sharpe. 1972. "Modal Matching Applied to a Discontinuity in Planar Surface Waveguide." *Electron. Lett.* 8: 28-29.

Clarricoats, P.J.B., and K. R. Slinn. 1967. "Numerical Solution of Waveguide Discontinuity Problems.", *Proc. IEE.* 114: 878.

Collin, R. E. 1991. *Field Theory of Guided Wave.* New York: IEEE Press.

Gordon, E., I. 1988. "High Power Single Mode Lasers" *United States Patent.* No. 4783.788.

Hahn,W.C. 1941. "A New Method for the Calculation of Cavity Resonator." *J. Appl. Phys.*12:62-68.

Haus, H.A. 1984. *Waves and Fields.* Prentice Hall.

Jackson, J. D. *Classical Electrodynamics, Second Edition.* New York: Wiley.

Lewin, L. 1975. *Theory of Waveguides.* New York: Wiley.

Lorrain, P., D. P.Corson, and F. Lorrain. 1988. *Electromagnetic Fields and Waves.* W.H. Freeman.

BIBLIOGRAPHY
(Continued)

Marcuvitz, N. 1965. *Waveguide Handbook.* New York: Dover.

Miller, A.R. 1988. *Fortran Programs for Scientists and Engineers.* SYBEX.

Owyang, G. H. 1981. *Foundation of Optical Waveguides.* New York: Elsevier.

Sander, K.F., and G. A. L. Reed. 1978. *Transmission and Propagation of Electromagnetic Waves.* Cambridge University Press.

Snyder, A. W. and J. D. Love. 1983. *Optical Waveguide Theory.* New York: Chapman & Hall.

Streetman, B.G. 1990. *Solid State Electronic Devices.* Practice Hall.

Tamir, T. and A. A. Oliner. 1963. "Guided Complex Waves,Part I." *Proc. IEE.* 110: 310.

Tamir, T. and A. A. Oliner. 1963. "Guided Complex Waves,Part II." *Proc. IEE.* 110: 325.

Thompsom, G.H.B. 1980. *Physics of Semiconductor Laser Devices.* New York: Wiley.